# SPIM Thèse de Doctorat

 école doctorale sciences pour l'ingénieur et microtechniques

 UNIVERSITÉ

 UNIVERSITÉ

 DEABOURT

### Systèmes de vision à grande dynamique auto-adaptables

Mustapha Bouderbane





# SPIM Thèse de Doctorat

École doctorale sciences pour l'ingénieur et microtechniques UNIVERSITE DE FOURCEGENE N° 0 2 - 0 1 - 2 0 1 5

THÈSE présentée par

### MUSTAPHA BOUDERBANE

pour obtenir le Grade de Docteur de l'Université de Bourgogne

Spécialité : Electronique

# Systèmes de vision à grande dynamique auto-adaptables

Unité de Recherche : Laboratoire Électronique, Informatique et Image

Soutenue publiquement le 23 novembre 2020 devant le Jury composé de :

Fredéric MORAIN-NICOLIER	Président du jury	Professeur à l'Université de Reims			
Francois BERRY	Rapporteur	Professeur à l'Université de			
		Clermont Auvergne			
Abdelhafid EL OUARDI	Rapporteur	Maître de Conférences HDR à			
		l'Université de Paris Saclay			
Dominique GINHAC	Directeur de thèse	Professeur à l'Université de			
		bourgogne			
JULIEN DUBOIS	Codirecteur	Maître de conférences HDR à			
		l'Université de bourgogne			
BARTHÉLÉMY HEYRMAN	Coencadrant	Maître de conférences à l'Université			
		de bourgogne			

"La vraie science est une ignorance qui se sait."

Michel de Montaigne

### REMERCIEMENTS

Les travaux présentés dans ce manuscrit ont été réalisés au sein du laboratoire LE2I dans l'équipe de recherche architecture des systèmes de vision. Je tiens à remercier vivement les deux rapporteurs qui m'ont fait l'honneur d'accepter de juger ce travail : le Professeur Francois Berry (Professeur à l'Université de Clermont Auvergne) et le Maître de conférences Abdelhafid El Ouardi (Maître de conférences à l'Université de Paris Sud). Je remercie les professeurs Fabrice Mériaudeau et Dominique Ginhac, directeurs du laboratoire LE2I au cours de ma thèse, de m'avoir accueilli dans leur laboratoire. Je remercie Julien Dubois, Maître de conférences au laboratoire LE2I et l'Université de Bourgogne et Franche-Comté et codirecteur de ma thèse pour ses relectures attentives de ce manuscrit. Merci également pour ton coaching qui a commencé dès le début de ma thèse. Je te remercie également pour ta disponibilité, ton bureau est toujours resté ouvert pour répondre à mes nombreuses et diverses questions que se soit en liaison directe avec ma thèse ou administratives. Merci pour tes interventions régulières chez les différents doctorants où tu en profitais pour donner un petit commentaire toujours très constructif tant sur le côté technique de la manipulation en préparation ou la démarche scientifique à suivre. Si cette thèse s'est si bien passée, c'est aussi parce que tu y es pour beaucoup. Je remercie mon encadrant Barthélémy Heyrman, Maître de conférences à l'Université de Bourgogne et Franche-Comté, pour son accompagnement le long de ma thèse et son aide précieux surtout dans la partie du développement informatique et également pour ses relectures du manuscrit. Je remercie mon directeur de thèse. le Professeur Dominique Ginhac, Professeur à l'université de Bourgogne et Franche-Comté de m'avoir fait confiance dans mes démarches en me laissant la liberté de choix dans mes recherches, tout en étant disponible à mes questionnements durant mes années de thèse.

Mes remerciements s'adresse ensuite aux autres membres passés ou présents de Laboratoire LE2I équipe architecture des systèmes de vision (pole 5) : Fan Yang et Franck Marzani pour les discussions enrichissantes pendant les poses café. Je tiens à remercier tous les doctorants, ingénieurs et stagiaires à I3M pour les moments conviviales qu'on a pu partager et les moments de détente ping pong. Comme je remercie le Professeur El-Bay Bourennane, Professeur à l'Université de Bourgogne et Franche-Comté et directeur adjoint de l'école doctorale EDSPIM, pour ses conseils et son aide dans les enseignements que j'ai assuré avec lui dans le département informatique et électronique de l'Université de Bourgogne et Franche-Comté. Mes remerciements sont destinés aussi au personnels administratif de laboratoire LE2I surtout les secrétaires Dounia Radi, Melanie Arnoult et Dorielle-Angès Batis pour leur aide dans tout ce qui est démarches administratives. Je tiens à exprimer ma profonde gratitude, à ma famille surtout mon petit frère surnommé Bara et mes deux chères amis Massinissa et Abderrazak pour leur soutien inconditionnel. Je termine par rendre grâce pour celle qui a cru en moi dès les premiers jours de ma thèse, ma femme Sajida, qui n'a jamais cessé de me motiver et de m'encourager, merci infiniment pour ta patience et ta sagesse dans la gestion de mon tempérament dans les périodes de stress

## TABLE DES MATIÈRES

Re	emerc	iemen	ts		vii		
Та	ble d	es mat	ières		ix		
1	intro	troduction					
2	Ľim	agerie	agerie à grande gamme dynamique				
	2.1	Introdu	uction		6		
		2.1.1	La radio	métrie	7		
			2.1.1.1	L'effet du système optique de la camera sur la radiance .	8		
		2.1.2	La photo	ométrie	9		
		2.1.3	Colorime	étrie	10		
		2.1.4	La gamr	ne dynamique	14		
	2.2 La génération d'image HDR						
		2.2.1	les capte	eurs d'image conventionnels	15		
		2.2.2 Propriétés des capteurs d'image conventionnels					
			2.2.2.1	La sensibilité	17		
			2.2.2.2	La gamme dynamique <i>DR</i>	17		
			2.2.2.3	Le mode de capture	18		
			2.2.2.4	La vitesse de la lecture	19		
			2.2.2.5	La conversion analogique/numérique	19		
		2.2.3	Les capt	eur d'images HDR	20		
			2.2.3.1	Les capteur d'images HDR à un seul temps d'intégration .	20		
			2.2.3.2	Les capteurs d'images HDR à multiples temps d'intégration	21		
			2.2.3.3	Les capteur d'images HDR logarithmiques	21		
			2.2.3.4	Performances et limitations des capteurs HDR	22		
	2.3	Créatio	on de cor	tenu HDR à partir de multiples expositions LDR	23		
	2.3.1 La génération HDR dans le domaine de la radiance						

			2.3.1.1	Acquisition de multiples expositions	24
			2.3.1.2	Récupération de la courbe de réponse inverse	24
			2.3.1.3	Construction de la carte de radiance	28
			2.3.1.4	Tone mapping	30
			2.3.1.5	Format d'image HDR	30
	2.4	Conclu	usion		31
3	Ľart	efact d	e ghost o	en imagerie HDR	33
	3.1	Introdu	uction		35
		3.1.1	La défini	tion de ghost	36
	3.2	Détect	ion de gh	ost	36
		3.2.1	Détection	n et suppression du mouvement de la caméra	37
	3.3	Détect	ion et sup	pression du mouvement des objets dans la scène	39
		3.3.1	Méthode	s de correction de ghost sans utilisation d'image de référence	39
			3.3.1.1	Détection basée sur le gradient	39
		3.3.2	Méthode	s de correction utilisant une seule image de référence	41
			3.3.2.1	détection basée sur la variance	41
			3.3.2.2	Détection basée sur l'entropie	42
			3.3.2.3	Détection basée sur un seuillage multi-niveaux	44
			3.3.2.4	Détection basée sur un bitmap	45
			3.3.2.5	Détection basée sur RANSAC	46
			3.3.2.6	Détection basée sur des densités de probabilités conjointes	47
			3.3.2.7	Detection basée sur la compensation du mouvement	49
			3.3.2.8	Détection de ghost par le calcul du flux optique	50
			3.3.2.9	Détection et suppression de ghost par une fonction de pondération	52
		3.3.3	Méthode	es de détection utilisant plusieurs images de références	52
			3.3.3.1	Détection basée sur l'ordre du pixel	52
			3.3.3.2	Détection basée sur la prédiction	53
		3.3.4	La suppr	ession du ghost pour la vidéo HDR	54
			3.3.4.1	Méthode basée sur l'estimation de mouvement par bloc	54
			3.3.4.2	Méthode utilisant un algorithme d'optimisation basé sur un patch temporellement cohérent	55
	3.4	Systèr	nes maté	riels pour la vidéo HDR	58

\_\_\_\_\_

		3.4.1	Système basé sur l'adaptation de la gamme dynamique	58
		3.4.2	Système HDR basé sur plusieurs capteurs d'image	60
	3.5	Artéfa	cts générés par la suppression de ghost	61
		3.5.1	Métriques pour la suppression de ghost	62
			3.5.1.1 Métrique de fusion	62
			3.5.1.2 Métrique de cohérence de gradient	63
			3.5.1.3 Métrique de différence visuelle	64
			3.5.1.4 Métrique de gamme dynamique	64
			3.5.1.5 Métrique unifiée de la qualité de la suppression de ghost	65
	3.6	Conclu	usion	66
4	Éva imp	luation lantatio	l logicielle d'algorithmes de suppression de ghost en vue d'une on matérielle	e 67
	4.1	introdu	uction	68
	4.2	Le log	jiciel GhostEva	68
	4.3	Test et	t évaluations des méthodes de deghosting	71
		4.3.1	Évaluation de la méthode basée sur l'ordre des pixels	72
		4.3.2	Évaluation de la méthode basée sur la MTB	75
		4.3.3	Évaluation de la méthode basée sur la prédiction	80
		4.3.4	Évaluation de la méthode basée sur la variance	84
		4.3.5	Évaluation de la méthode basée sur la fonction de pondération simple originale	87
		4.3.6	Évaluation de la méthode basée sur la fonction de pondération simple Améliorée	89
	4.4	Compa	araison des résultats des différentes méthodes de suppression de	
		ghost		91
		4.4.1	Comparaison des résultats obtenus pour la scène du bureau	92
		4.4.2	Comparaison des résultats obtenu pour la scène du jardin	93
		4.4.3	Comparaison des résultats obtenu pour la scène du cycliste	93
		4.4.4	Comparaison des résultats moyens des trois scènes	95
	4.5	Conclu	usion	96
5	Imp	lantatio	on matérielle de la suppression de ghost et son intégration à la אר	00
			union	100
	5.1	mirodi	uciion	100

	5.2	La car	néra HDF	3	100			
		5.2.1	Le gesti	onnaire de la mémoire vive MMU	103			
		5.2.2	Le contr	ôleur de l'interface Ethernet	106			
		5.2.3	Le modu	Ile de tone mapping	109			
		5.2.4	La génération HDR					
	5.3	Le mo	dule de la	a suppression de ghost	111			
		5.3.1	Suppres	sion de ghost par la méthode basée sur l'ordre des pixels .	111			
			5.3.1.1	Résultats de l'implantation globale utilisant l'ordre des pixels	113			
		5.3.2	Suppres	sion de ghost par la fonction de pondération	115			
			5.3.2.1	Implantation matérielle de la fonction de pondération utili- sant le DL de l'exponentiel	116			
			5.3.2.2	L'efficacité de la génération HDR	117			
			5.3.2.3	L'efficacité en suppression de ghost	118			
			5.3.2.4	Implantation matérielle de la fonction de pondération amé- liorée utilisant le DL de l'exponentiel	119			
			5.3.2.5	Le module de la génération HDR utilisant le module de la suppression de ghost	122			
			5.3.2.6	Résultats d'acquisition	125			
		5.3.3	Suppres	sion de ghost par la fonction de pondération tabulée	126			
			5.3.3.1	Le module de la génération HDR utilisant le module de suppression de ghost tabulé	127			
		5.3.4	Résultat bulée de	s obtenus par notre caméra HDR utilisant l'architecture ta- module de suppression de ghost	131			
			5.3.4.1	Efficacité en génération HDR	132			
			5.3.4.2	Efficacité en préservation de texture	134			
			5.3.4.3	Efficacité en génération HDR et suppression de ghost	134			
	5.4	Conclu	usion		136			
6	Con	clusio	n et pers	pectives	139			
	6.1	Perspe	ectives .		141			
GI	ossa	ire			143			
Bi	bliog	raphie			145			

Та	ble des figures	157
Lis	ste des tables	167
I	Annexes	173
A	La base de donnée d'images LDR utilisée dans les tests logiciels	175
В	Les menus déroulant de logiciel GhostEva	177
С	La carte imprimé de capteur d'image	181

## 1 INTRODUCTION

Les systèmes d'acquisition d'image et de la vidéo (précisément le domaine des caméras et des appareils photo) sont en évolution permanente. Ceci s'est d'autant plus accéléré depuis le passage des systèmes d'acquisition à base de pellicule photographique (film) aux caméras numériques qui utilisent des capteurs d'images électroniques (CCD ou CMOS). Désormais, l'évolution des systèmes d'acquisition d'image et de vidéo suit la même loi d'évolution que ces circuits intégrés, permettant de convertir le signal lumineux en une grandeur électrique mesurable. L'augmentation de l'intégration et la miniaturisation des circuits microélectroniques assurent une évolution spontanée dans l'amélioration de la définition des capteurs d'image. L'augmentation de la définition des capteurs d'image a été pendant longtemps la caractéristique dont la concurrence était la plus concentrée, car la technologie le permet et le rendement est meilleur que le rendement des implémentations logicielles d'augmentation de la définition (résolution). Des plus grandes définitions sont alors obtenues par l'augmentation de l'échantillonnage spatial de signal lumineux capturé. Cette augmentation d'échantillonnage est assurée par l'utilisation d'un nombre élevé de pixels (de nos jours, il existe des capteurs d'image dont la définition est de quelques dizaines de millions de pixels) sur la matrice de capteur d'image donc des images capturées d'une haute définition et qui contiennent de plus en plus de détails. Cependant, en analysant des images de grande définition issues de ces systèmes, on s'aperçoit que l'augmentation du nombre de pixels ne récupère pas forcément plus de détails dans l'image de la scène capturée, car des blocs de plusieurs pixels ont souvent des valeurs identigues. Ces valeurs identigues sont liées aux échantillons du signal lumineux qui ont des niveaux d'intensité proches par zones et les valeurs des intensités lumineuses sont très éloignées entre une zone et une autre. Si le capteur utilise un pas de quantification grand afin de couvrir la dynamique la plus large possible de signal lumineux, la différence entre les intensités lumineuses dans les zones proches est plus faible que le pas de quantification, ce qui donne des pixels avec des valeurs identiques. Pour récupérer plus de détails dans ces zones, il faut donc augmenter la résolution ou la dynamique du pixel afin qu'il puisse représenter l'information le plus fidèlement possible en la quantifiant sur un nombre plus élevé de niveaux et avec un petit pas de quantification. L'image obtenue dont les pixels sont d'une résolution élevée (codés sur un mot

excédant les *8bits*) est dite de grande gamme de dynamique ou simplement HDR (HDR acronyme anglais de High Dynamic Range), ouvrant un nouveau domaine de concurrence pour les fabricants des systèmes d'acquisition d'image et une nouvelle branche dans le domaine de l'imagerie nommée imagerie à grande gamme de dynamique. Cette nouvelle branche d'imagerie est prometteuse (voire nécessaire) pour différents domaines d'application. Elle permet de fournir des images d'une grande qualité : riches en détails donc en informations.

On constate l'intérêt, voire même la nécessité de l'utilisation de l'imagerie à grande gamme de dynamique dans divers domaines :

- a-La vidéosurveillance : La sécurité dans les magasins, sociétés, routes, banques et aéroports, et d'autres lieux, est assurée en partie ou exclusivement (dans les horaires non travaillés) par la vidéosurveillance. La vidéosurveillance est très contraignante la nuit où les scènes observées par les caméras de surveillance sont illuminées par des projecteurs. Une personne malveillante assise dans l'ombre ou derrière une source de lumière est difficilement reconnaissable par les caméras conventionnelles. On retrouve les mêmes problèmes avec les plaques d'immatriculation des voitures lorsque les phares de cette dernière sot allumés en face de l'objectif de la caméra.
- b-Les systèmes autonomes de conduite : Prenant l'exemple de l'entrée et la sortie d'un tunnel, les caméras utilisées dans les systèmes de décision des voitures autonomes récupèrent des images totalement saturées blanches de la sortie du tunnel ou des images complètement noires de l'entrée du tunnel (car la différence d'exposition entre l'intérieure et l'extérieur de tunnel est très grande), ce qui fait que la décision dans ces circonstances sera prise en se basant sur des informations incomplètes, voire erronées.
- c-Le médical : L'imagerie médicale est un recours obligatoire dans le diagnostic de certaines maladies. Bien que les systèmes d'imagerie médicale (scanner, IRM, écho, etc.) soient différents de ceux de l'imagerie grand public, l'imagerie à grande gamme de dynamique peut être très utile dans ce domaine. Elle permet par exemple d'augmenter le contraste entre les tissus cancéreux et les tissus sains dans le cas de diagnostic du cancer.
- d-L'astrophysique : Les objets astrophysiques ont des intensités lumineuses très écartées. Prenant l'exemple de l'éclipse solaire, la surface de la lune capturée par une caméra conventionnelle est totalement noire, car l'intensité de la lumière du soleil situé derrière est beaucoup plus élevée que celle réfléchie de la surface de la lune. De même dans le cas de la photographie d'une nébuleuse, si on essaye d'avoir des détails du nuage de la poussière, le cœur de cette dernière se sature complètement en blanc. Il est impossible d'obtenir simultanément les détails du nuage et du cœur de la nébuleuse.

Les images HDR peuvent être obtenues par deux différents procédés. Dans le premier, les images HDR sont générées de manière séparée ou isolée de système d'acquisition (le système d'acquisition dans ce cas est traditionnel ou de faible gamme de dynamique LDR : acronyme anglais de Low Dynamic Range) en utilisant l'une des techniques de la génération HDR (qu'on trouve dans l'état de l'art) implémentée en logiciel (la technique utilisée reproduit des images HDR à partir d'images LDR, c'est l'une des solutions les plus adaptées et les moins coûteuses, permettant d'augmenter la gamme dynamique des images acquises par les capteurs d'image conventionnels). Le deuxième procédé consiste à obtenir les images HDR directement à partir du système d'acquisition. Ce dernier peut intégrer un capteur d'image permettant des acquisitions HDR natives (très contraignant : coût élevé, verrou technologique et peu fiable) ou garder l'utilisation d'un capteur LDR et générer des images HDR par reconstruction en post traitement au sein du système d'acquisition par l'intégration d'une des techniques de la génération HDR à partir d'images LDR. Bien que plusieurs algorithmes de la génération HDR à partir d'images LDR sont facilement intégrables au sein d'un système acquisition, atteindre certaines performances (faible latence, cadence images élevée) s'avèrent difficiles car les algorithmes doivent être adaptés à l'architecture matérielle du système et aux conditions de l'acquisition (prise en compte du mouvement dans la scène ou de la caméra).

Le projet de cette thèse s'inscrit dans le développement d'une caméra HDR (système d'acquisition d'images et vidéo HDR par une reconstruction HDR intégrée au sein du système). Cette dernière implantera de nouvelles méthodes d'acquisition, de traitement d'images à grande gamme de dynamique. Pour obtenir la plus grande gamme dynamique de la scène à photographier ou à filmer, la caméra proposée utilise la technique d'acquisition successive de plusieurs images de la scène en faisant varier le temps d'exposition. Cet ensemble d'images est ensuite combiné afin de produire l'image finale à grande dynamique en utilisant l'algorithme proposé par Debevec et Malik [Debevec et al., 2008]. Bien que de nombreuses solutions algorithmiques aient été développées en particulier dans le cadre d'applications grand public de photo numérique, il existe peu de travaux à l'heure actuelle sur la conception matérielle de systèmes de vision dédiés capables de produire en temps réel du contenu HDR. La mise au point d'un tel système matériel et logiciel implique donc de mettre en place une démarche d'adéquation algorithme architecture permettant d'adapter à la fois la chaîne de traitement d'images et le système matériel d'acquisition et de traitement, permettant ainsi de tirer parti des possibilités d'acquisition de contenu HDR. Cette thèse est une continuité de travaux de recherche déjà menés au sein du laboratoire LE2I, qui ont permis de développer un premier prototype de caméra à grande gamme de dynamique. Le prototype développé utilise 3 images LDR pour générer une HDR à une cadence de 60 images/s. Cependant, le prototype initial intègre seulement l'algorithme original de Debevec et Malik [Debevec et al., 2008] sans aucune adaptation aux scènes dynamiques (contenant du mouvement), sachant que cet algorithme est proposé pour des scènes statiques uniquement, son utilisation

pour des scènes dynamique génère inévitablement des artefacts de fantômes appelés "ghost" dans les images HDR. La plus grande partie des travaux de cette thèse sont dédiés à l'étude et au développement de plusieurs algorithmes pour la correction des effets de ghost qui peuvent apparaître dans les images HDR regénérées à partir d'une pile d'images LDR à différents temps d'exposition. L'objectif final et ambitieux de cette thèse est de développer une nouvelle génération de systèmes de vision surclassant les systèmes actuels et capables de fournir en temps réel du contenu HDR. Les enjeux et les retombées de ces travaux sont multiples sur le plan logiciel et matériel, en termes d'acquisition et de stockage d'images, en termes de traitement en temps réel du flot vidéo d'images et en termes de restitution d'images de grande qualité.

À l'issue de cette partie introduisant le contexte de recherche mené dans le cadre de cette thèse et les objectifs tracés, le chapitre 2 sera consacré à l'état de l'art de la génération HDR, où on citera les méthodes principales permettant la création d'images HDR (pour la photo comme pour la vidéo) tout en fournissant leurs avantages et inconvénients. On détaillera aussi dans ce chapitre les différents systèmes matériels avec leurs contraintes et verrous.

Le deuxième chapitre sera dédié aux artefacts du mouvement (ghost), qui accompagnent la génération HDR utilisant la technique de multiples expositions. Il détaillera les méthodes de la suppression et de la correction de cet artefact. Une sélection de quelques méthodes de correction de ghost sera faite à l'issue de ce chapitre pour une étude plus poussée en termes d'efficacité de correction et d'adaptation à une éventuelle implantation matérielle.

Dans le chapitre 3, on étudiera l'efficacité des méthodes de suppression de ghost sélectionnées. L'étude est basée sur les résultats obtenus en implémentant ces méthodes en logiciel. L'étude réalisée sur ces méthodes déterminera l'algorithme ou les algorithmes à implanter en matériel en se basant sur une comparaison des différents résultats des méthodes étudiées et ceci de manière quantitative (en utilisant des métriques reconnues par la communauté scientifique de l'imagerie HDR) et aussi de manière subjective selon le rendu visuel des différentes méthodes.

Le quatrième et dernier chapitre sera consacré à la présentation complète de la caméra HDR proposée et sa mise en œuvre. Le système d'acquisition (et de vision) HDR sera présenté de manière exhaustive. On détaillera son architecture matérielle dont on donnera différentes versions selon la méthode de suppression de ghost implantée. Des résultats d'implantation matérielle seront donnés pour les différentes versions de l'architecture matérielle globale de la caméra ainsi que pour les différents modules principaux qui la constituent. Les résultats d'acquisitions obtenus en utilisant la caméra HDR seront également présentés et commentés pour chacune des versions de l'architecture matérielle globale.

2

### L'IMAGERIE À GRANDE GAMME DYNAMIQUE

#### Sommaire

2.1	Introd	luction					
	2.1.1	La radio	métrie	7			
		2.1.1.1	L'effet du système optique de la camera sur la radiance	8			
	2.1.2	La photo	ométrie	9			
	2.1.3	Colorime	étrie	10			
	2.1.4	La gamr	ne dynamique	14			
2.2	La gé	nération	n d'image HDR				
	2.2.1	les capte	eurs d'image conventionnels	15			
	2.2.2	Propriét	és des capteurs d'image conventionnels	17			
		2.2.2.1	La sensibilité	17			
		2.2.2.2	La gamme dynamique DR	17			
		2.2.2.3	Le mode de capture	18			
		2.2.2.4	La vitesse de la lecture	19			
		2.2.2.5	La conversion analogique/numérique	19			
	2.2.3	Les capt	teur d'images HDR	20			
		2.2.3.1	Les capteur d'images HDR à un seul temps d'intégration	20			
		2.2.3.2	Les capteurs d'images HDR à multiples temps d'inté- gration	21			
		2.2.3.3	Les capteur d'images HDR logarithmiques	21			
		2.2.3.4	Performances et limitations des capteurs HDR	22			
2.3	Créati	on de co	ntenu HDR à partir de multiples expositions LDR	23			
	2.3.1	La géné	ration HDR dans le domaine de la radiance	23			
		2.3.1.1	Acquisition de multiples expositions	24			
		2.3.1.2	Récupération de la courbe de réponse inverse	24			
		2.3.1.3	Construction de la carte de radiance	28			
		2.3.1.4	Tone mapping	30			
		2.3.1.5	Format d'image HDR	30			
2.4	Concl	usion		31			

#### 2.1/ INTRODUCTION

Les images capturées par la majorité des caméras numériques de nos jours sont de faible gamme dynamique. Les pixels de ces images sont généralement codés sur 8 bits ou 24 bits pour les images couleurs (8 bits par canal de la couleur rouge, vert et bleu). Sachant que le nombre maximal de niveaux qui peuvent être représentés sous 8 bits est seulement de 256, cela reste insuffisant pour représenter la gamme dynamique d'une scène réelle qui peut contenir des zones moins éclairées (sombres) et des zones fortement éclairées avec une grande différence d'intensité de la luminance entre ces différentes zones. Par exemple, la luminance du soleil est de 10<sup>5</sup> et un objet situé dans l'ombre a une luminance inférieure à  $10^2$ . Ainsi dans ce cas typique, 256 niveaux d'intensité ne suffisent pas pour représenter l'objet dans l'ombre et le soleil en même temps. L'imagerie à grande gamme dynamique HDRI (acronyme anglais de High Dynamic Range Imaging) est une technique introduite afin de remédier à cette limitation. Il existe deux grandes classes de génération d'images à grande gamme dynamique. La première classe cherche à introduire des améliorations au niveau du capteur d'image afin d'obtenir un capteur d'image nativement de grande gamme dynamique. Dans ce cas, l'image HDR est générée directement par le capteur [ni et al., 2012]. Cependant, le coût du système d'acquisition HDR est très important puisque il nécessite la conception et la fabrication d'un capteur spécifique. La seconde classe maintient l'utilisation d'un ou plusieurs capteurs à faible gamme dynamique. Dans le cas d'utilisation de plusieurs capteurs, les images à différents temps d'expositions sont capturées simultanément à l'aide de différents capteurs visualisant la même scène [Griffiths et al., 2017]. Chacun des capteurs est configuré avec un temps d'intégration différent. Alors que dans le cas de l'utilisation d'un capteur conventionnel unique, ce dernier est configuré pour acquérir une succession d'images à différents temps d'expositions, les images étant capturées séquentiellement l'une après l'autre à des instants différents (avec un espacement dans le temps). Ensuite, l'image HDR est construite par un post-traitement des différentes images acquises. La construction HDR peut se faire à partir d'une implantation logicielle ou matérielle, réalisé soit dans le domaine de l'image [Tom Mertens, 2007], soit dans le domaine de la radiance [Debevec et al., 2008]. Enfin, il existe également une alternative pour la génération d'image HDR. Elle réside dans une opération d'étalement de dynamique, réalisée en post acquisition d'une seule image à faible gamme dynamique capturée par un capteur conventionnel [Banterle et al., 2009].

Ce premier chapitre a pour vocation de fournir les notions fondamentales de l'imagerie à grande gamme dynamique. Nous commencerons par donner les notions de base qui définissent cette branche d'imagerie, tout en montrant les limitations de l'imagerie conventionnelle, ensuite nous détaillerons les différentes techniques de la génération HDR mono et multi-capteurs, y compris les avantages et les inconvénients de chacune de ces techniques.

#### 2.1.1/ LA RADIOMÉTRIE

Le terme radiométrie désigne la branche de la physique mesurant le rayonnement électromagnétique (ondes radio, rayons X, rayons  $\gamma$ , lumière visible par l'œil humain, etc). La lumière se caractérise par deux propriétés, une propriété corpusculaire donnée par le photon et une propriété ondulatoire donnée par l'onde électromagnétique dont l'énergie est mesurée en *Joules* qu'on note  $Q_e$ . La lumière se propage dans les milieux dits transparents comme l'air, l'eau, le verre et l'espace de manière rectiligne. Il est intéressant de considérer les quantités dérivées pour quantifier la manière dont la lumière se propage. Les quantités dérivées incluent la mesure de l'énergie de rayonnement par unité de temps, d'espace ou d'angle. La puissance électromagnétique est la mesure du flux de rayonnement électromagnétique par unité de temps donnée par  $\Phi_e = \frac{dQ_e}{dt}$ . L'intensité électromagnétique représente la mesure de la puissance par unité d'angle solide définie par  $I_e = \frac{d\Phi_e}{d\omega}$ .  $\omega$  est l'angle solide en stéradian définie par  $d\omega = \frac{dS}{R^2}$  qui est le quotient de la zone sur une sphère de rayon *R* de la projection de la courbe sur la sphère et  $R^2$ . Cette valeur devient  $d\omega = \frac{dS}{R^2}$ . cos  $\theta$  lorsque la surface est inclinée d'un angle  $\theta$  par rapport à la direction d'observation comme illustré sur la figure 2.1.





La radiance (luminance énergétique)  $L_e$  quantifie la puissance électromagnétique émise/reçue par une surface émettrice/réflective dans une direction particulière. Elle est mesurée en Watt par stéradian par mètre carré ( $W.sr^{-1}.m^{-2}$ ). La radiance  $L_e$  est donnée par :

$$L_e = \frac{d^2 \Phi_e}{d\omega dA \cos \theta},\tag{2.1}$$

Où  $\Phi_e$  est la puissance du flux électromagnétique en Watt (*W*), *A* est la surface apparente en mettre carré ( $m^2$ ),  $\omega$  est l'angle solide et  $\theta$  est l'angle fait par la perpendiculaire à la surface de vision avec la direction comme illustré sur la figure 2.2.



FIGURE 2.2 – Radiance : une énergie incidente sur une surface unitaire dA à partir d'un ensemble d'unités de directions  $d\omega$  [Reinhard et al., 2010].

#### 2.1.1.1/ L'EFFET DU SYSTÈME OPTIQUE DE LA CAMERA SUR LA RADIANCE

Les systèmes d'objectif des caméras sont généralement construits à partir d'une série de lentilles sphériques individuelles. Elles sont fabriquées en verre ou en plastique. Elle partagent avec le diaphragme un axe commun, l'axe optique. Le diaphragme contient des éléments opaques avec une ouverture approximativement circulaire pour permettre le passage de la lumière. Les éléments opaques de diaphragme sont ceux qui limitent le plus la propagation angulaire du faisceau de rayons. Quant l'image est prise, l'obturateur est ouvert pour une courte durée, pendant lequel la lumière incidente se focalise sur le capteur d'image en utilisant les lentilles d'objectif. Cette utilisation de lentilles limites le nombre de directions à partir desquels la lumière est reçue. La focalisation de la puissance électromagnétique sur une surface en utilisant une lentille est définie par l'irradiance (éclairement énergétique noté  $E_e$  en  $W/m^2$ ). L'irradiance observée au niveau d'un pixel de capteur d'image (ou d'une image optique donnée) est proportionnelle à la radiance du point correspondant de la scène. La relation linéaire entre la radiance et l'irradiance dans le cas de lentilles minces est donnée par la formule ci-après :

$$E = L\frac{\pi}{4} \left(\frac{d}{f}\right)^2 \cos \alpha^4, \tag{2.2}$$

où d, f sont respectivement le diamètre et la distance focale de la lentille, et  $\alpha$  l'angle que fait l'axe optique de la lentille avec le faisceau de lumière incidente. L'influence de

l'angle  $\alpha$  sur la proportionnalité entre la radiance et l'irradiance est corrigée lors de la construction des lentilles actuelles.

#### 2.1.2/ LA PHOTOMÉTRIE

Les surfaces reflètent la lumière en modifiant éventuellement sa composition spectrale. Ainsi, la lumière réfléchie transmet l'information spectrale à la fois de la source de lumière qui éclaire un point de surface et le facteur de réflexion (la réflectance) de la surface à ce point. L'œil humain est sensible aux longueurs d'ondes situées dans l'intervalle [380nm – 830nm] (figure 2.3 ) ce qui montre l'incapacité de ce dernier de visualiser des longueurs d'ondes à l'extérieur de cet intervalle. L'œil humain n'est pas aussi sensible à toutes les longueurs d'ondes. Il est plus sensible aux couleurs verte et jaune (longueur d'onde proche de 555nm) et sa sensibilité décroît en fur à mesure qu'on s'éloigne de cette longueur d'onde dans les deux sens de l'intervalle. En outre, il existe des différences de



FIGURE 2.3 – Le spectre de la lumière visible.

sensibilité à la composition spectrale de la lumière chez les individus. Cependant, cette gamme de sensibilité est suffisamment petite pour que la sensibilité spectrale d'un observateur humain avec une vision normale puisse être représentée avec une seule courbe. Cette courbe est normalisée par la Commission Internationale de l'éclairage (CIE) et est connue sous le nom de courbe d'efficacité lumineuse photopique de CIE (ou simplement courbe  $V(\lambda)$ ). La figure 2.4 montre le tracé de cette courbe.

Étant donné que l'on s'intéresse à la perception de la lumière par l'œil humain, la composition spectrale peut être pondérée selon  $V(\lambda)$ . La photométrie est la science de la mesure de la lumière dans des unités pondérées. Toutes les grandeurs radiométriques présentées précédemment ont des équivalents photométriques, qui sont listées dans le tableau 2.1. En pondérant spectralement des quantités radiométriques et en utilisant la réponse spectrale de l'œil humain  $V(\lambda)$ , les grandeurs radiométriques sont converties en quantités photométriques. La puissance lumineuse est une puissance électromagnétique pondérée par photométrie. Elle est mesurée en lumens, qui est défini comme 1/683W de



FIGURE 2.4 – Le tracé de la courbe d'efficacité lumineuse photopique donnée par le CIE[Reinhard et al., 2010].

puissance rayonnante à une fréquence de  $54 \times 10^{13} Hz$ . Cette fréquence correspond à la longueur d'onde pour laquelle l'œil humain est le plus sensible (environ 555nm). Si le flux lumineux est mesuré sur un angle solide élémentaire, la quantité obtenue est l'intensité lumineuse donnée en lumen par stéradian. Un lumen par stéradian correspond à un candela. L'illuminance et l'éclairement sont donnés en lumens par mètre carré, alors que la luminance est spécifiée en candela par mètre carré. La luminance est une quantité perçue. C'est la radiance pondérée par photométrie. Elle constitue une mesure approximative de la façon dont une surface brillante apparaît. La luminance est l'unité photométrique la plus pertinente pour l'imagerie à grande gamme dynamique. La pondération spectrale de la radiance est calculée en multipliant chaque composante spectrale par la valeur correspondante donnée par la fonction de pondération, suivi par une intégration de résultat sur l'intervalle des longueurs d'ondes du spectre visible de l'œil humain.

$$L_{\nu} = \int_{a380}^{830} L_{e,\lambda} V(\lambda) d\lambda$$
(2.3)

#### 2.1.3/ COLORIMÉTRIE

La colorimétrie est une filière d'imagerie qui a comme objectif de caractériser numériquement les couleurs en établissant un lien entre la perception de la couleur par l'œil humain et la description physique de la lumière. La perception de la couleur par l'œil humain est déterminée par trois types de cônes (L, M et S). La sensibilité de ces trois différents cônes change selon les longueurs d'ondes. Les cônes S ont une forte sensibi-

Grandeurs radior	nétriques	
Grandeur	Unité	notation
Énergie électromagnétique	Joule	$Q_e$
Puissance électromagnétique	$W\left(J.s^{-1} ight)$	$\Phi_e$
Intensité électromagnétique	$W.sr^{-1}$	$I_e$
Radiance	$W.sr^{-1}.m^{-2}$	$L_e$
Irradiance	$W.m^{-2}$	$E_e$
Grandeurs photor	nétriques	
Grandeur	Unité	notation
Quantité de lumière	lm.s	$Q_v$
Puissance lumineuse	lumen	$\Phi_v$
Intensité lumineuse	candela	$I_{v}$
Luminance	$cd.m^{-2}$	$L_v$
Illuminance	$lux (lm.m^{-2})$	$E_{v}$

Tableau 2.1 – Correspondance entre les grandeurs radiométriques et celle photométriques

lité aux ondes à 420 nm de longueur (Bleue), les cônes M sont sensibles aux longueurs d'onde de 530 nm (Vert) et les cônes L sont sensibles aux longueurs d'ondes à 560 nm (Rouge). La lumière dans le spectre visible est en fait une variable multidimensionnelle, où chaque dimension est associée à une longueur d'onde particulière. Cependant, la couleur visible est une projection de cette variable multidimensionnelle basées sur les trois couleurs primaires, correspondant aux trois types de cônes. Chaque type de cône présente une sensibilité à une région du spectre (couleur). Les cônes plus sensibles à la lumière bleue (CS), d'autres à la lumière verte (CM) et le troisième type de cônes est plus sensible à la lumière rouge (CL). Cette projection peut être décrite mathématiquement sous forme d'un produit de la distribution de puissance spectrale, la réponse spectrale du type de cônes ( $CL(\lambda), CM(\lambda)$  et  $CS(\lambda)$ ) et  $\Phi(\lambda)$  qui est la réponse spectrale des bâtonnets.

$$R = \int_{\lambda} \Phi(\lambda) C_L(\lambda) d\lambda$$
 (2.4)

$$V = \int_{\lambda} \Phi(\lambda) C_M(\lambda) d\lambda$$
 (2.5)

$$B = \int_{\lambda} \Phi(\lambda) C_S(\lambda) d\lambda$$
 (2.6)

Le codage de la couleur dans le système de vision humaine (SVH) est tridimensionnel ce qui limite le nombre de couleurs distinguables. En outre, deux stimuli de différentes distributions de puissance spectrale peuvent être considérés comme ayant la même couleur si seulement leurs projections R, V et B correspondent. Cette propriété du SVH s'appelle métamérisme. Afin de décrire de manière unique la gamme des couleurs visibles, le CIE a standardisé en 1931 un ensemble de primaires pour l'observateur colorimétrique standard. Étant donné que les sensibilités spectrales du cône n'étaient pas encore connues, les primaires étaient basées sur une expérience de correspondance des couleurs, dans laquelle les stimuli monochromatiques de longueur d'onde particulière correspondaient à un mélange des trois primaires monochromatiques (435.6 nm, 546.1 nm et 700 nm). Les valeurs du mélange d'équivalent de la couleur des primaires pour chaque longueur d'onde ont donné les primaires R, V et B illustrés dans la figure 2.5.Le R primaire a une valeur négative ce qui est considéré comme un inconvénient de cette procédure, car la partie négative représente des gammes de la couleur très saturée qui ne peuvent être représentées dans une gamme visible ou une gamme physiquement réalisable. Pour associer ces couleurs à la gamme valide, la saturation des couleurs est corrigée en ajoutant une certaine quantité de lumière monochromatique qui induit une augmentation des valeurs de toutes les composantes primaires R, V et B pour qu'elles deviennent positives.



FIGURE 2.5 – Fonctions de correspondance des couleurs pour les stimuli R, V et B de CIE et l'observateur standard 2°[Myszkowski et al., 2008].

Dans la continuité de la définition des couleurs primaires négatives, le CIE a introduit les primaires X, Y, Z dont la figure 2.6 montre leurs courbes. La primaire Y représente la luminance, sa fonction de correspondance de la couleur correspond à la fonction d'efficacité lumineuse (la fonction est montrée dans la figure 2.4). Ces nouvelles primaires sont utilisées comme référence dans les formules de conversion des couleurs.

Les coordonnées de chromaticité *x*, *y* sont souvent utilisées pour représenter pratiquement la couleur dans une représentation bidimensionnelle. Mais ces coordonnées doivent être accompagnées de la valeur de la luminance Y pour décrire pleinement la couleur.

$$x = \frac{X}{X + Y + Z} \tag{2.7}$$

$$y = \frac{Y}{X + Y + Z} \tag{2.8}$$



FIGURE 2.6 – Fonctions de correspondance des couleurs pour les stimuli X, Y et Z de CIE et l'observateur standard 2°[Myszkowski et al., 2008].

Les coordonnées de chromaticité x et y ne décrivent pas efficacement les différences visibles entre les couleurs. Ce qui a mené le CIE à définir les balances de chromacité uniformes afin de donner une meilleure représentation des différences perceptuelles de la couleur. Cependant les balances de chromacité uniforme n'intègrent pas le niveau de la luminance dans leur couleur ce qui est considéré comme une limitation majeure, car la différence de la couleur peut être fortement dépendante du niveau de la luminance réelle. Afin de résoudre ce problème le CIE a introduit deux espaces uniformes de la couleur  $L^*a^*b^*$  et  $L^*u^*v^*$ .

La fonction L\* des deux espace est la même :

$$L^* = 116(Y/Y_n)^{1/3} - 16$$
 (2.9)

Les fonction des composantes  $a^*, b^*, u^*, v^*$  des deux espaces sont définies comme suit :

$$a^* = 500[(X/Y_n)^{1/3} - (Y/Y_n)^{1/3}]$$
(2.10)

$$b^* = 500[(Y/Y_n)^{1/3} - (Z/Y_n)^{1/3}]$$
(2.11)

$$u^* = 13L^*(u' - u'_n) \tag{2.12}$$

$$v^* = 13L^*(v' - v'_n) \tag{2.13}$$

où les composantes données en indice *n* indiquent la couleur du blanc de référence, les composantes  $u' = \frac{4X}{X+15Y+3Z}$  et  $v' = \frac{9Y}{X+15Y+3Z}$  sont les valeurs des balances uniformes de chromacité définies par le CIE 1976.

#### 2.1.4/ LA GAMME DYNAMIQUE

En principe, le terme gamme dynamique est utilisé dans l'ingénierie pour définir le ratio entre la plus grande et la plus petite quantité à l'étude. Prenant en considération la terminologie dans le domaine de l'image, la quantité physique observée est le niveau de luminance. Il existe plusieurs mesures de gamme dynamique. Elles dépendent de l'application (le tableau 2.2 contient un résumé des différentes mesures de gamme dynamique). Le ratio de contraste est une mesure utilisée dans les systèmes d'affichage (écrans). Elle définit le ratio de la luminance de la couleur la plus brillante (représentant un point blanc sur l'afficheur ou l'écran) que l'écran peut produire et la luminance de la couleur la plus sombre (représentant le noir). Lorsque un écran n'émet pas de la lumière (au niveau zéro), le premier niveau contrôlable au dessus de zéro est considéré comme le niveau le plus sombre définissant le noir, et cela dans le but d'éviter les formes de résultat indéterminé en divisant par zéro. Le ratio dans le cas des écrans est souvent normalisé pour que la deuxième valeur soit toujours 1 (100 : 1 au lieu de 10 : 0.1). La gamme logarithmique d'une exposition est adoptée par l'imagerie à grande gamme dynamique afin de mesurer la gamme dynamique des scènes à photographier ou à filmer. Dans ce cas, la gamme est mesurée entre la luminance la plus brillante et la plus sombre d'une scène donnée. La gamme est calculée en réalisant le logarithme base 10 du point le plus brillant et du plus sombre. L'utilisation des valeurs logarithmiques décrit fidèlement la différence perçue en gamme dynamique contrairement au ratio de contraste. La latitude de l'exposition est définie comme la gamme de la luminance qu'un film peut capturer moins la gamme de la luminance de la scène photographiée. Son expression utilise un logarithme base 2. Ce type de formalisme est adapté à l'échelle des paramètres de l'exposition (fermeture d'ouverture ainsi appelée f-stops et vitesse d'obturateur) où un pas dans ces paramètres double ou divise par 2 la quantité de lumière capturée. Dans notre cas, pour les caméras numériques, on utilise souvent le ratio signal sur bruit (SNR acronyme anglais de Signal to Noise Ratio). Il est mesuré comme le quotient de l'intensité de signal qui commence à saturer le capteur d'image et l'intensité de signal minimale qui peut être observée audessus de niveau de bruit de capteur d'image. Le SNR est représenté en décibels [dB] en utilisant le logarithme de base 10 multiplié par un facteur 20.

Nom	La fonction	Exemple	Contexte
Le ratio de contraste	$CR = \frac{Y_{max}}{Y_{bruit}} : 1$	500:1	écran
Log de la gamme d'exposition	$D = log_{10}(\frac{Y_{max}}{Y_{hruit}})$	2.7 ordres	Image HDR
	$D = log_2(\frac{Y_{max}}{Y_{hruit}})$	9stops	Photographie
Pique de ratio signal sur bruit	$D = 20 \times log_{10}(\frac{Y_{max}}{Y_{hruit}})$	53[dB]	Capteur d'image

Tableau 2.2 – La mesure des gammes dynamiques et leurs expression selon le contexte. [Miller et al., 2013]

#### 2.2/ LA GÉNÉRATION D'IMAGE HDR

La génération de contenu d'une grande gamme dynamique peut se faire de deux manières différentes. La première méthode utilise un capteur d'image de grande gamme dynamique, conçu spécialement pour capturer le maximum de la gamme dynamique de la scène à filmer. La deuxième méthode maintient l'utilisation d'un capteur d'image de faible gamme dynamique (capteur conventionnel) et utilise différentes techniques afin de créer des images HDR à partir de contenu LDR issus de ce type de capteur.

#### 2.2.1/ LES CAPTEURS D'IMAGE CONVENTIONNELS

Un capteur d'image est un élément principal dans la composition d'une camera ou un appareil photo numérique. Le capteur d'image est composé essentiellement d'une matrice bidimensionnelle de pixels qui convertissent la lumière incidente sur leurs surfaces en signaux électriques. L'image couleur est crée en déposant un filtre couleur (CFA acronyme anglais de Color Filter Array, comme par exemple, le filtre couleur classique de Bayer) sur la matrice des pixels de capteur d'image. En utilisant ce filtre, chaque pixel produit un signal électrique analogique qui correspond à une composante des trois couleurs du filtre (rouge, vert ou bleu) [El Gamal et al., 2005]. Les signaux électriques analogiques issus des pixels sont lus à la sortie de la matrice des pixels et numérisés par un convertisseur analogique numérique (CAN).

Il existe deux grandes familles de capteurs d'image, les capteurs CCDs (acronyme anglais de Charge Coupled Devices) et les capteurs CMOS (acronyme anglais de Complementary Metal-Oxide-Semiconductor). Les deux familles de capteurs d'images utilisent des photo sites (pixels ou photo détecteurs) à base de Métal-Oxyde-Semi-conducteur (MOS). Ces capteurs accumulent le signal (charges électriques) dans chacun des pixels proportionnellement à l'intensité de l'illumination locale, ce qui assure un échantillonnage spatiale[Litwiller, 2001] pour une durée de temps dite temps d'exposition ou d'intégration. Ensuite dans le cas des CCDs, les charges accumulées par les pixels sont transférées d'une manière séquentielle à une structure de sortie commune qui convertit les charges en tensions électriques. Celles ci sont stockées de manière temporaire avant d'être envoyées à l'extérieur comme illustré sur la figure 2.7.



FIGURE 2.7 – Le capteur d'image CCD et sa liaison au PCB de la caméra.

Dans le cas des capteurs d'images CMOS, la conversion des charges électriques en tension est réalisée au niveau pixel (voir la figure 2.8). Cette différence dans la technique de la lecture a un impact important sur l'architecture de capteur d'image et ses capacités (lecture individuelle de pixel, région d'intérêt, amplification au niveau pixel, etc). Néanmoins, elle introduit quelques limitations (sensibilité plus faible que celle des CCDs, bruit électronique, etc). Cependant, de nouvelles technologies de fabrication d'imageurs CMOS, notamment la technologie 3D et la technologie illumination en face arrière (BSI acronyme anglais de Back Side Illumination) permettent de minimiser ces limitations.



FIGURE 2.8 – Le capteur d'image CMOS et sa liaison au PCB de la caméra.

#### 2.2.2/ PROPRIÉTÉS DES CAPTEURS D'IMAGE CONVENTIONNELS

Dans cette partie, nous détaillerons les propriétés et caractéristiques des capteurs conventionnels influant leurs gammes dynamiques de manière directe (gamme dynamique, sensibilité) ou indirecte dans le cas où le capteur conventionnel est utilisé pour générer des images de haute gamme dynamique (mode de capteur, rapidité, conversion analogique/numérique, fenêtrage ou région d'intérêt ROI).

#### 2.2.2.1/ LA SENSIBILITÉ

La sensibilité d'un capteur d'image est définie comme le minimum d'illuminance détectable ou bien le nombre de photons incidents nécessaires pour passer d'un niveau de gris à un autre. La sensibilité est dépendante de plusieurs facteurs dont on cite les deux plus importants :

- Le facteur de remplissage : Il est défini comme le rapport de la surface active d'un pixel sur la taille de ce dernier. Il détermine sa capacité à accueillir les photons incidents et à les convertir en charges électriques. Plus le facteur de remplissage est grand pour une taille d'un pixel donné, plus ce dernier collecte des photons, ce qui se traduit par une meilleure sensibilité. La sensibilité peut être améliorée en augmentant la taille d'un pixel à faible facteur de remplissage mais cette option limite la résolution spatiale du capteur d'image et elle augmente le coût de fabrication de la puce.
- l'efficacité quantique QE : Appelé aussi rendement quantique, elle correspond à la fraction du flux des photons qui contribue au photocourant dans un pixel. L'augmentation de l'efficacité quantique améliore le rapport signal sur bruit et la gamme dynamique du capteur d'image[Fowler et al., 1998].

Une comparaison réalisée entre les ratios de l'efficacité quantique QE et le bruit de l'obscurité ( $N_D$ ) de chacun des deux types de capteurs CMOS et CCD [KAMAKURA, 2017], a montré que la sensibilité des derniers capteurs d'images CMOS dépasse nettement celle de leurs homologues CDDs.

#### 2.2.2.2/ LA GAMME DYNAMIQUE DR

La gamme dynamique d'un capteur d'image est donnée par le rapport entre le niveau de saturation d'un pixel et son signal de seuil (le ratio entre la valeur maximale  $I_{max}$  et la valeur minimale  $I_{min}$  de courant électrique généré par le pixel). La valeur minimale de courant électrique correspond à la plus faible valeur de courant électrique détectable

après la sortie de niveau de bruit directement. La gamme dynamique d'un capteur est donnée par l'équation 2.14.

$$D = 20 \times \log_{10}(\frac{I_{max}}{I_{min}})$$
(2.14)

Dans des circonstances comparables, les capteurs d'images CCD dépassent les capteurs CMOS d'un facteur de deux [Litwiller, 2001], ce constat reste valable jusqu'en 2011. A partir de cette année et jusqu'en 2015 les performances (gamme de dynamique, vitesse, etc) des capteurs CMOS ont largement dépassé celles des CCD avec l'amélioration de la technologie des capteurs CMOS en BSI (illumination de la face arrière) et l'émergence de la technologie 3D dans la fabrication des imageurs CMOS [KAMAKURA, 2017].

#### 2.2.2.3/ LE MODE DE CAPTURE

Le mode de capture correspond à la possibilité d'arrêter l'intégration arbitrairement. On distingue deux modes de capture : le mode de capture instantané (en anglais Global Shutter) et le mode de capture en déroulement (en anglais Rolling Shutter). Le premier mode consiste à interrompre l'intégration des photons de manière simultanée pour tous les pixels de l'image. Les valeurs des pixels sont alors conservées jusqu'à leur lecture. Ce mode de capture est implémenté dans le cas où le capteur doit faire des acquisitions d'images dans des conditions particulières (scène dynamique avec des objets qui se déplacent à une grande vitesse, lorsque le temps d'acquisition est court, acquisitions d'images synchronisées ). Les deux types de capteurs CCD et CMOS sont dotés de ce mode de capture. Le deuxième mode de capture l'image ligne par ligne, le temps d'exposition est le même pour toutes les lignes. Le temps qui sépare entre les différentes intégrations est aussi constant. Le fait que l'intégration des différents lignes constituant l'image ne démarre pas au même moment introduit des déformations prévisibles pour les objets en mouvement dans la scène pendant une pause, ces déformations sont connues sous le nom de "jello effect". Le mode de lecture en déroulement n'est supporté que par les capteurs d'image CMOS, les capteurs CCDs ne peuvent bénéficier de ce mode. La capture en mode rolling shutter, permet de lire une ligne pendant que les autres lignes continuent l'intégration de photons, ce qui améliore le SNR. Néanmoins, ce mode de capture est limité par l'artefact dit "effet jello".

#### 2.2.2.4/ LA VITESSE DE LA LECTURE

Les capteurs CMOS sont significativement meilleurs que leurs homologues CCDs en terme de rapidité. Ceci est lié directement à l'architecture des capteurs d'image CMOS qui permet l'implémentation de la majorité des fonctionnalités de la caméra (conversion photons/charges électriques, amplification, circuits de lecture, conversion analogique/numérique, générateur d'horloge, etc) à l'intérieur du capteur. Cette intégration au sein de la puce du capteur minimise les délais de propagation des signaux électriques et en conséquence l'augmentation de la vitesse d'acquisition.

#### 2.2.2.5/ LA CONVERSION ANALOGIQUE/NUMÉRIQUE

Le convertisseur analogique/numérique CAN est un élément important dans une caméra. Il est l'élément en charge de la conversion des signaux analogiques provenant des pixels de capteur d'image sous forme d'une tension électrique en valeurs numériques manipulables par système numérique à base de DSP, CPU, GPU ou FPGA. Dans le cas des capteurs d'image CMOS, le CAN est implémenté au sein du capteur (selon l'architecture, un ou plusieurs CAN peuvent être utilisés) [Sony, 2008] alors que dans le cas des capteurs CCD la conversion analogique/numérique se fait à l'extérieur de la puce de capteur d'image. Les capteurs CMOS conventionnels (LDR) intègrent généralement des CAN d'une résolution allant de 8 à 12 bits . Le CAN est considéré comme un des points clés dans la fabrication des capteurs CMOS. Différents types de CAN existent (flash, sigma-delta, approximations successives, etc). Le choix d'un type donné de CAN est réalisé afin de favoriser une propriété donnée de capteur (convertisseur de type flash offre une conversion rapide mais, il prend plus de surface en silicium. Le CAN sigmadelta utilise moins de surface en silicium mais sa fréquence de conversion est basse, etc). Cependant, le défi réside dans l'augmentation de la résolution du convertisseur car la résolution de celui-ci détermine la gamme dynamique finale de capteur. L'augmentation de la résolution de CAN doit être réalisée tout en gardant un taux d'échantillonnage du signal analogique élevé permettant la capture en temps réel de la lumière incidente et sa transmission numérique au processeur de traitement de la caméra. L'augmentation de la résolution de CAN dépend de plusieurs paramètres (consommation, rapidité de conversion, surface en silicium) et un compromis adéquat doit être trouvé pour obtenir un CAN adapté au système visé.

#### 2.2.3/ LES CAPTEUR D'IMAGES HDR

Comme mentionné précédemment, la gamme dynamique des capteurs d'images conventionnels est très limitée (généralement elle ne dépasse pas les 62*dB* pour des données codé sur 10 bits). La méthode la plus simple réside dans la conception d'une caméra à grande gamme dynamique en remplaçant le capteur d'image de faible gamme dynamique par un autre d'une grande gamme dynamique. Le changement de capteur d'images est accompagné par d'autres modifications de système de base, mais ces modifications restent simples, comme une modification de l'interface de contrôle du nouveau capteur et un élargissement du bus de données pour supporter les données de pixels codés sur un plus grand nombre de bits. On trouve plusieurs architectures proposées dans l'état de l'art permettant de concevoir des capteurs d'image à grande gamme dynamique.

#### 2.2.3.1/ LES CAPTEUR D'IMAGES HDR À UN SEUL TEMPS D'INTÉGRATION

Dans le cas des capteurs HDR n'utilisant qu'un seul temps d'intégration, l'augmentation de la gamme dynamique se fait de deux manières différentes. La première, consiste à augmenter la capacité effective d'intégration de charge, et par conséquent le nombre de charges accumulées. Le courant électrique maximum est ainsi amplifié lorsque le temps d'intégration augmente [Kang et al., 2005]. La deuxième méthode tend à diminuer le bruit, pour détecter des courants très petits. Mais cette dernière souffre de la saturation des pixels lors de l'utilisation de temps d'intégration long. On trouve différentes architectures dans l'état de l'art qui cherchent soit à éviter qu'un pixel sature par l'ajout d'une capacité au niveau du pixel, ce qui lui permet d'accumuler l'excès de charges lorsqu'un pixel sature [Akahane et al., 2006b] [Sugawa et al., 2005]. D'autres architectures cherchent à exploiter la saturation des pixels par différentes méthodes. Une méthode présentée par [Miyagawa et al., 1995] utilise un pixel possédant deux photo diodes de différentes tailles, la petite diode est utilisée comme détecteur de la saturation et activatrice de la deuxième photo diode. En plus de l'insertion de cette seconde photo diode de plus petite taille, cette méthode nécessite la mémorisation du temps d'intégration qui a induit la saturation afin de reconstruire l'information d'illumination à la fin de l'acquisition. Une autre méthode basée sur la définition d'un pixel se réinitialisant en limite de saturation est proposée par [Leñero-Bardallo et al., 2017]. A chaque initialisation de pixel un compteur calculant le nombre d'initialisations est incrémenté [McIlrath, 2001] [Rhee et al., 2003] et la valeur finale de pixel est obtenue à partir de la dernière quantité de charge accumulée et du compteur d'initialisation.

#### 2.2.3.2/ LES CAPTEURS D'IMAGES HDR À MULTIPLES TEMPS D'INTÉGRATION

L'architecture des capteurs d'images HDR à multiples temps d'intégration est basée sur la lecture des pixels plusieurs fois à différents temps d'exposition, toute en les effaçant à chaque lecture. L'image finale est reconstruite à partir des informations recueillies pendent les différentes lectures. Cette méthode implique une fréquence image basse en raison de la lecture de la matrice des pixels plusieurs fois afin d'obtenir une seule image [Schrey et al., 2002]. Pour maintenir une fréquence proche de celles des capteurs conventionnels une architecture intégrant deux sorties par pixels et deux circuits d'amplification par colonne est proposée par [Yadid-Pecht et al., 2003]. Dans une autre architecture proposée par [Hosticka et al., 2003], le capteur présente des pixels qui peuvent être lus et initialisés plusieurs fois pour constituer une seule image. Le choix, le stockage et le codage du signal optimum sont réalisés en fonction de temps d'intégration de ce dernier. Néanmoins l'architecture proposée est complexe, et nécessite l'usage d'une mémoire externe.

Ce type d'architectures est difficile à mettre en place. Cette difficulté est due à la complexité algorithmique (la décision sur le nombre de lectures par pixel, le calcul de la dynamique optimum souhaitée, la fréquence images, etc) et la complexité de son électronique car elle nécessite plus d'électronique au sein d'une même puce afin d'intégrer les multiples lectures des différents pixels pour une seule image. Par conséquent, d'autres contraintes s'ajoutent à ce type de capteurs, citant le bruit de lecture qui n'est plus uniforme (car le nombre de lectures n'est pas le même pour tous les pixels) et aussi un facteur de remplissage faible car le circuit de lecture de ces capteur est beaucoup plus grand que celui des capteurs conventionnels ce qui réduit la sensibilité de ces derniers.

#### 2.2.3.3/ Les capteur d'images HDR logarithmiques

Ce type de capteurs exploite la combinaison de deux réponses, une linéaire et l'autre logarithmique appelés capteurs lin-log. La société Photonfocus a implémenté et breveté un capteur lin-log intégrant un seul transistor additionnel par pixel par rapport au pixel des capteurs conventionnels. Le transistor ajouté agit sur le niveau d'illumination au sein de chaque pixel. Le pixel maintient une réponse linéaire et une sensibilité stable pour une illumination locale faible qui ne dépasse pas une certaine valeur définie comme seuil, mais dès que l'illumination locale dépasse ce seuil, la réponse du pixel devienne progressivement logarithmique. Une architecture de pixel lin-log proposée par [Hara et al., 2005] et[Storm et al., 2006] fournit deux informations (un signal issu de l'intégration en utilisant la réponse linaire et le signal issu de la réponse logarithmique successivement).

Référence	Dynamique en dB	Taille de pixel <i>um</i> <sup>2</sup>	Facteur de remplissage en %	Taille de capteur proposé	Frame rate en fps
[Akahane et al., 2006b]	100	$7.5 \times 7.5$	ND	$640 \times 480$	30
[Akahane et al., 2006a]	207	$20 \times 20$	ND	$70 \times 68$	ND
[Posch et al., 2011]	143	$30 \times 30$	10	$304 \times 240$	30
[Leñero-Bardallo et al., 2	105 – 130	$25 \times 25$	10	96 × 128	0.5 - 200
[McIIrath, 2001]	100	$30 \times 30$	ND	$64 \times 64$	ND
[Rhee et al., 2003]	120	19 × 19	50	$320 \times 256$	1k
[Schrey et al., 2002]	90	$10 \times 10$	41	1024×1024	15
[Yadid-Pecht et al., 2003]	108	$20.4 \times 20.4$	15	$64 \times 64$	100(kpixel/s)
[Hosticka et al., 2003]	120	$26 \times 26$	65	$1025 \times 1024$	50
[Hara et al., 2005]	124	$7.5 \times 7.5$	37	$640 \times 480$	ND
[Storm et al., 2006]	120	$5.6 \times 5.6$	33	$352 \times 288$	26

Tableau 2.3 – Les performances et caractéristiques clés des capteurs HDR proposés dans l'état de l'art.

Une autre architecture proposée par [Hara et al., 2005] recombine les deux signaux issus de la réponse logarithmique et celle linaire en un seul signal final au niveau du pixel. [Storm et al., 2006] présente un capteur d'image qui implémente un algorithme permettant le choix de meilleur signal. Cet algorithme sélectionne le signal linéaire lorsque le pixel est non saturé sinon il utilise le signal issu de la réponse logarithmique.

Bien que la gamme de dynamique de ces capteurs soit élevée (plus que 120*dB*), ils présentent plusieurs faiblesses, un ratio signal sur bruit faible, un bruit de modèle fixe (FPN acronyme anglais pour Fixed Pattern Noise) élevé et une faible dynamique pour les faibles illuminations.

#### 2.2.3.4/ Performances et limitations des capteurs HDR

Les performances des différentes architectures proposées afin de permettre la capture d'une grande gamme dynamique, sont données dans le tableau 2.3.

Les architectures proposées pour la conception des capteurs d'images HDR sont d'une complexité extrême, et les capteurs d'image résultants représentent plusieurs limitations :

- Sensibilité réduite à cause du faible facteur de remplissage.
- Bruit de lecture amplifié imposé par l'utilisation de plusieurs transistors au niveau pixel.
- Limitation de la vitesse de lecture liée au mode d'intégration et au mode de lecture imposés par l'architecture choisie.
- Résolution spatiale des capteurs très petite au regard de la grande taille du pixel
HDR.

Coût de fabrication de capteur élevé.

Les limitations des différentes architectures de capteur d'image nativement HDR présentées précédemment n'ont pas permis à ces capteurs de passer à l'étape de fabrication à grande échelle. Leur fabrication se limite à des applications spécifiques et coûteuses. L'indisponibilité de ces capteurs pour des applications grand public a mené les acteurs de ce secteur à utiliser une alternative moins chère, permettant la génération HDR à partir de multiples expositions issues d'un capteur LDR.

# 2.3/ CRÉATION DE CONTENU HDR À PARTIR DE MULTIPLES EXPO-SITIONS LDR

C'est une solution bas coût et flexible permettant la création d'image HDR à partir d'une liste d'images LDR capturées en utilisant différent temps d'exposition comme illustré par la figure 2.9. La dynamique désirée peut être paramétrée ou ajustée selon le besoin en agissant sur le nombre d'images LDR à utiliser pour créer une seule image HDR. Pour acquérir la gamme dynamique complète d'une scène, il est possible de prendre une liste idéale d'images LDR allant d'une image sous exposée (complètement noir) à une image surexposée (saturée blanche) avec un pas d'exposition adapté. Cependant, l'utilisation d'un grand nombre d'images LDR peut compromettre les performances du système final (fréquence d'images, résolution et complexité de l'implantation logiciel/matérielle). En principe, il existe deux méthodes qui permettent la génération HDR en exploitant des images issues d'un ou de plusieurs capteurs d'image LDR. La première méthode est réalisée dans le domaine de l'image, elle génère des image pseudo-HDR (HDR-like en anglais) [Raman et al., 2007]. La deuxième méthode consiste à fusionner les différentes expositions LDR dans le domaine de la radiance. Dans cette rubrique, on ne détaille que la génération d'image HDR dans le domaine de la radiance, car ce mode offre la possibilité de recouvrir la gamme dynamique réelle de la scène capturée, ce qui ne peut être réalisé en utilisant le mode de génération dans le domaine de l'image qui génère des images contenant plus de détails mais de faible gamme dynamique.

## 2.3.1/ LA GÉNÉRATION HDR DANS LE DOMAINE DE LA RADIANCE

La création d'images HDR par la fusion de multiples expositions LDR dans le domaine de la radiance est basée sur la récupération de la carte de radiance. Cette dernière est construite en utilisant la courbe inverse de la réponse de système d'acquisition. La



FIGURE 2.9 – Création d'image HDR à partir de multiples expositions LDR.

courbe inverse sert à convertir les pixels des images LDR en valeurs de radiance en les préparant à l'étape de fusion de données en utilisant un algorithme de création HDR.

## 2.3.1.1/ ACQUISITION DE MULTIPLES EXPOSITIONS

L'acquisition des différentes expositions (images LDRs à différents temps d'expositions) peut se faire de deux manières différents. La première utilise plusieurs capteurs d'images LDR visualisant une même scène soit avec des angles de vue différentes soit en utilisant des diviseurs de faisceau (miroirs semi-réfléchissants) comme le montre la figure 2.10. La seconde méthode utilise un seul capteur d'image LDR, les images LDR à différents temps d'exposition sont capturées d'une manière séquentielle en faisant varier le temps d'intégration de capteur à chaque acquisition.



FIGURE 2.10 – La visualisation d'une seule scène avec plusieurs capteurs d'images via un diviseur de faisceau[Griffiths et al., 2017].

## 2.3.1.2/ RÉCUPÉRATION DE LA COURBE DE RÉPONSE INVERSE

La fonction de réponse inverse (courbe de réponse inverse)  $(f^{-1}())$  d'une caméra ou un appareil photo est primordiale afin de remonter à l'information de l'irradiance. Cette

fonction dépend du capteur d'image du système d'acquisition et de son optique (système de lentilles) ce qui fait que chaque camera ou appareil photo transforme la lumière incidente différemment des autres en appliquant sa fonction de transfert (f()). La fonction f() d'une caméra est non linéaire car contrairement au système de lentilles qui atténue la lumière incidente linéairement, le capteur d'image a une réponse non linéaire. La fonction f() peut être obtenue directement par une combinaison de la fonction de transfert du système optique et celle de capteur d'image utilisé, mais cette dernière généralement n'est pas donnée par le fabriquant car souvent elle fait l'objet d'une propriété intellectuelle. La fonction de transfert f() de système d'acquisition donne la relation qui lie la lumière incidente sur le capteur d'image en passant par le système de lentilles et les valeurs des pixels dans l'image résultante comme le montre la figure 2.11. Après la récupération de



FIGURE 2.11 – La fonction de transfert d'un système d'acquisition d'image.

la fonction f() et l'admission que cette fonction est strictement croissante et continue, on peut déduire la fonction inverse  $f^{-1}()$  à utiliser pour récupérer l'irradiance E correspondante à chaque pixel. La figure 2.12 montre le graphe d'une fonction de réponse inverse 2.15 estimée en utilisant l'algorithme de Debevec et Malik [Debevec et al., 1997].

$$g(Z_{ij}^{p}) = \ln E_{ij} + \ln \Delta t_{ij}^{p}, \qquad (2.15)$$

Où *i*, *j* sont les coordonnées du pixel *Z*, et  $\Delta t^p$  est le temps d'exposition de l'image *p*.

On trouve plusieurs méthodes de récupération de la courbe de réponse de capteur. On a sélectionné les quatre méthodes les plus connues dans le domaine de HDR. Chacune des fonctions de réponse de capteur est associée à une fonction de pondération permettant la fusion de données.

1. Méthode de Mann : [Mann et al., 1995]



FIGURE 2.12 – Courbe de réponse inverse g(Z), pour la base de donnée d'images de mémoriel de Debevec [Debevec et al., 2008].

C'est la toute première méthode proposée en 1995 par Mann et al. Elle permet de fusionner les intensités lumineuses réelles de plusieurs expositions en une seule image HDR où la valeur de pixel HDR est flottante. La courbe de réponse est reconstruite automatiquement à partir d'une liste d'images LDR à différentes expositions. Ensuite la courbe estimée est utilisée en combinaison avec une fonction de pondération pour fusionner les données. La fonction de pondération utilisée est la dérivée de la fonction de réponse du système  $f'^{-1}$ .

2. Méthode de Mitsunaga et Nayar : [Mitsunaga et al., 1999]

Cette méthode est proposée par Mitsunaga et Nayar. Ils calculent une fonction de réponse radiométrique par une estimation à l'aide d'un polynôme d'ordre élevé. L'estimation de la courbe de réponse est réalisée sans avoir recours aux temps d'exposition utilisés lors de la capture des différentes images LDR. Les temps d'exposition sont aussi estimés et sont affinés lors de processus d'estimation de la courbe de réponse. Néanmoins la méthode est limitée par le calcul d'un polynôme d'un certain ordre et des coefficients sélectionnés. Les auteurs ont émis l'hypothèse que la courbe de réponse soit continue. Pour la fusion de donnée, Mitsunaga et Nayar utilisent une fonction de pondération basée sur la théorie du signal de telle sorte que plus la valeur de pixel est élevée moins cette valeur est influencée par le bruit du fond qui est supposé être constant. Une amélioration de cette méthode est donnée par Ward [Reinhard et al., 2005], en multipliant la courbe de réponse obtenue par une fonction "chapeau" large, afin de réduire l'effet imprévisible que peuvent avoir les valeurs de pixels proches de saturation en blanc ou en noir.

Méthode de Debevec et Malik : [Debevec et al., 2008]
 Debevec et Malik [Debevec et al., 2008], [Debevec et al., 1997] ont proposé un al-

gorithme qui relie la valeur du pixel Z et le rapport de l'irradiance E et du temps d'exposition  $\Delta t$  par la fonction de réponse f() donnée par l'équation 2.16 :

$$Z_{ij}^p = f(E_{ij}\Delta t_p). \tag{2.16}$$

La méthode de Debevec et Malik est basée sur le fait que la valeur de l'exposition reste la même si on divise la valeur de l'irradiance *E* et on double le temps d'exposition  $\Delta t$ . La courbe de réponse est déduite par la résolution d'un ensemble d'équations linéaires par la décomposition en valeurs singulières (SVD) tout en supposant que la courbe de réponse est continue. Ils utilisent une fonction de pondération triangulaire simple de type "chapeau" donnée sur la figure 2.13 pour fusionner les valeurs d'irradiance récupérées à partir des différentes expositions disponibles. Leur fonction de pondération est basée sur l'hypothèse que les pixels de milieu de la plage de dynamique du capteur (proches de 128 dans le cas d'un capteur à sortie codée en 8 bits) sont les plus fiables et les mieux exposés.

4. Methode de Robertson : [Robertson et al., 2003] Robertson et al. [Robertson et al., 2003] ont proposé une méthode probabiliste pour estimer la courbe de réponse de leur système d'acquisition. La méthode proposée est basée sur un modèle d'observation. Elle est considérée non linéaire et elle est approximée par une spline cubique avec un nombre modéré de nœuds afin de faciliter le calcul de sa dérivée. La dérivée de la fonction de la réponse de la caméra est ensuite utilisée comme une fonction de pondération, en se basant sur le fait que la fiabilité des valeurs de pixel est en corrélation directe avec la sensibilité de l'appareil aux changements de la lumière. Les auteurs supposent que la sortie intègre un bruit sous forme d'une gaussienne. La fonction de pondération de Robertson et al. impose un facteur de pondération plus important aux pixels qui ont des temps d'exposition longs. L'algorithme proposé par Robertson et al. est reconnu pour sa réduction de bruit lors de la construction d'image HDR.

Une étude sur les méthodes de récupération de la fonction de réponse d'une caméra [Gençtav et al., 2011] montre que l'algorithme proposé par Debevec et Malik est le plus fiable. Cette étude a mesuré le nombre d'erreurs lors de la récupération de la fonction de réponse de la caméra (voir le tableau 2.4) pour chacune des méthodes.

On se basant sur les résultats montrés sur le tableau 2.4, la méthode de Debevec et Malik est choisie afin d'être intégré dans notre caméra, car elle donne le moins d'erreurs lors de la récupération de la CRF donc la plus fiable des deux autres méthodes.

Méthode	Erreur
[Debevec et al., 1997]	5538.77
[Mitsunaga et al., 1999]	11803.23
[Robertson et al., 2003]	6145.13

Tableau 2.4 – La comparaison de fiabilité entre les différentes méthode de récupération de la fonction de réponse de la caméra [Gençtav et al., 2011].

## 2.3.1.3/ CONSTRUCTION DE LA CARTE DE RADIANCE

Dès que la courbe de réponse de capteur est estimée par une des méthodes citées ci-dessus, son inverse  $f^{-1}()$  est utilisée pour récupérer les irradiances correspondantes aux différentes images prises à différents temps d'exposition  $\Delta t$ . Ensuite, les irradiances correspondantes aux pixels des images LDRs sont pondérées et fusionnées afin de reconstruire les valeurs réelles de la radiance de la scène. La pondération est utilisée pour sélectionner les pixels qui contribuent dans la reconstruction de l'irradiance finale en leur donnant plus du poids, et discriminer les pixels qui ne doivent pas contribuer significativement en leur imposant un poids faible lors de la fusion de données. La fonction générale permettant de calculer les valeurs d'irradiances en fusionnant les irradiances estimées de chacune des images LDR est donnée par l'équation 2.17 :

$$\ln E_{ij} = \frac{\sum_{p=1}^{P} \omega(Z_{ij}^{p}) [g(Z_{ij}^{p}) - \ln \Delta t_{p}]}{\sum_{p=1}^{P} \omega(Z_{ij}^{p})}.$$
(2.17)

où w(Z) est la fonction de pondération. Cette fonction est différente d'une méthode à l'autre selon le principe choisi par l'auteur de chacune des méthodes :

1. Mann [Mann et al., 1995] considèrent que la fiabilité de la valeur du pixel est en corrélation avec la sensibilité de l'appareil aux changements de lumière

$$\omega(Z) = \frac{1}{g'(Z))} \tag{2.18}$$

 Debevec [Debevec et al., 2008] s'appuient sur l'hypothèse "Plus le pixel est proche de milieu de plage plus sa valeur est fiable". La fonction de pondération est de forme triangulaire, la pointe au milieu représente le point le plus fort.

$$\omega(Z) = \begin{cases} Z \text{ pour } Z \le 127 \\ 255 - Z \text{ pour } Z > 127 \end{cases}$$
(2.19)

**3.** Mitsunaga [Mitsunaga et al., 1999] proposent une fonction de pondération basée sur le fait que l'augmentation de l'intensité du signal entraine une diminution de

bruit et par conséquent un rapport signal sur bruit élevé.

$$\omega(Z) = \frac{g(Z)}{g'(Z)}$$
(2.20)

4. Ward [Reinhard et al., 2005] proposent une fonction de pondération en multipliant la fonction de pondération de Mitsunaga par un filtre large de type "chapeau" afin de discriminer les pixels saturés.

$$\omega(Z) = \frac{g(Z)}{g'(Z)} \left( 1 - \left(\frac{Z}{127.5} - 1\right)^{12} \right)$$
(2.21)

 Robertson [Robertson et al., 2003] Considérant que la fiabilité des valeurs des pixels est en corrélation avec la sensibilité de l'appareil aux changements de lumière.

$$\omega(Z) = \frac{\Delta t^2}{g'(Z)}$$
(2.22)

La représentation graphique des différentes fonctions de pondération est donnée sur la figure 2.13



FIGURE 2.13 – Les fonctions de pondération normalisées associées aux courbes de réponse citées [Mann et al., 1995, Debevec et al., 1997, Mitsunaga et al., 1999, Reinhard et al., 2005].

Une étude approfondie sur les méthode de génération HDR [Yourganov et al., 2001], montre que la méthode de Debevec et Malik est la plus performante pour la vidéo, ce qui est lié directement à la simplicité de ses fonctions (CRF et fonction de pondération) et leurs robustesse. La fonction de pondération de Debevec et Malik est simple, elle détermine le poids en faisant une sélection selon la valeur de pixel directement sans autres calculs ni d'autres paramètres ce qui facilite l'intégration de cette fonction dans notre caméra basée sur un FPGA.

## 2.3.1.4/ TONE MAPPING

Le tone mapping est une technique de conversion d'images de grande gamme dynamique en images de faible gamme dynamique. Cette conversion est obligatoire pour visualiser une image HDR sur un écran LDR due à la non disponibilité des écrans permettant la visualisation directe d'un contenu HDR (Les écrans dits HDR de nous jours ne peuvent afficher un contraste au dessus de 1 : 65536 qui représente une dynamique de 96 *dB* alors que la gamme dynamique des images HDR peut dépasser largement 120 *dB*). Il existe deux grandes familles d'opérateurs de tone mapping :

- Tone mapping global : C'est un opérateur qui utilise une fonction spatialement uniforme [F. Drago et al., 2003], le calcul de la valeur tone mappée d'un pixel se fait en prenant en compte la valeur maximale et la valeur minimale de toute l'image HDR.
- 2. Tone mapping local : c'est un opérateur qui agit par zone. Contrairement à l'opérateur de tone mapping global ce dernier calcule la valeur de pixel tone mappé en prenant en compte la valeur maximale et la valeur minimale des pixel d'une zone de l'image HDR [Reinhard et al., 2002].

Les opérateurs de tone mapping global sont algorithmiquement plus simples par rapport aux opérateurs locaux, néanmoins ils ne préservent pas les détails lorsque la dynamique de l'image HDR est trop élevée.

#### 2.3.1.5/ FORMAT D'IMAGE HDR

Après la création d'image de grande gamme dynamique et afin de faciliter leurs manipulation (sauvegarde sous un support physique, transfert d'une machine à l'autre,...etc), trois différents formats sont présentés dans l'état de l'art (HDR, EXR et TIFF). Les formats les plus complets et flexibles sont le format EXR et le format TIFF. Ces deux formats permettent plusieurs encodages et plusieurs types de compression avec une possibilité de représenter des données HDR brutes sans compression ou avec compression. Les deux formats contient des champs de méta-données complets avec une possibilité d'insérer de nouveaux paramètres par l'utilisateur. Le tableau 2.5, donne les différents format HDR et leurs caractéristiques et points clés.

Format	Encodage	Compression	Méta-donnée	support
HDR	RGBE, XYZE	run-length	calibration, espace couleur, +à définir par l'utilisateur	open source soft- ware (Radiance)
TIFF	IEEE RGB LogLuv24 LogLuv32	pas de com- pression pas de com- pression run-length	calibration, espace couleur, +à définir par l'utilisateur	public domain library (libtiff)
EXR	Half RGB	wavelet, ZIP	calibration, espace couleur, fenêtrage, +à définir par l'utili- sateur	open source library (OpenEXR)

Tableau 2.5 – Les formats HDR[Reinhard et al., 2010].

## 2.4/ CONCLUSION

Les différentes sections de ce chapitre, ont permis de cerner le domaine d'imagerie à grande gamme dynamique. Dans un premier temps, les grandeurs physiques nécessaires pour l'étude de rayonnement électromagnétique ont été définies (radiance, irradiance, luminance, flux optique, etc). Ces grandeurs, sont utilisées dans le domaine de l'imagerie en général et dans le domaine de l'imagerie HDR en particulier. Les techniques principales permettant d'augmenter la dynamique capturée sont présentées comme une solution aux différentes limitations de l'imagerie conventionnelle. Parmi les différentes techniques de la création HDR, on a détaillé celle dite de multiples expositions qui consiste à utiliser des images LDR capturées en utilisant un capteur conventionnel et les fusionner dans le domaine de la radiance.

La technique de génération de HDR à partir de multiples expositions, est choisie pour les avantages qu'elle propose (gamme dynamique théoriquement infinie, faible coût de réalisation de système d'acquisition, etc). Son principal inconvénient provient de l'acquisition décalée dans le temps des différentes expositions. En cas de mouvement dans la scène, des artefacts apparaissent dans l'image HDR diminuant sa qualité. Ces artefacts appelés ghost seront détaillés dans le chapitre suivant.

# L'ARTEFACT DE GHOST EN IMAGERIE HDR

## Sommaire

3.1	Introd	troduction		
	3.1.1	La défin	ition de ghost	36
3.2	Détec	tion de g	host	36
	3.2.1	Détectio	n et suppression du mouvement de la caméra	37
3.3	Détec	tion et sı	uppression du mouvement des objets dans la scène.	39
	3.3.1	Méthode	es de correction de ghost sans utilisation d'image de ré-	
		férence		39
		3.3.1.1	Détection basée sur le gradient	39
	3.3.2 Méthodes de corr		es de correction utilisant une seule image de référence	41
		3.3.2.1	détection basée sur la variance	41
		3.3.2.2	Détection basée sur l'entropie	42
		3.3.2.3	Détection basée sur un seuillage multi-niveaux	44
		3.3.2.4	Détection basée sur un bitmap	45
		3.3.2.5	Détection basée sur RANSAC	46
		3.3.2.6	Détection basée sur des densités de probabilités conjointes	47
		3.3.2.7	Detection basée sur la compensation du mouvement	49
		3.3.2.8	Détection de ghost par le calcul du flux optique	50
		3.3.2.9	Détection et suppression de ghost par une fonction de	
			pondération	52
	3.3.3	Méthode	es de détection utilisant plusieurs images de références .	52
		3.3.3.1	Détection basée sur l'ordre du pixel	52
		3.3.3.2	Détection basée sur la prédiction	53
	3.3.4	La suppression du ghost pour la vidéo HDR		
		3.3.4.1	Méthode basée sur l'estimation de mouvement par bloc	54
		3.3.4.2	Méthode utilisant un algorithme d'optimisation basé sur	
			un patch temporellement cohérent	55
3.4	Systè	mes mate	ériels pour la vidéo HDR	58
	3.4.1	Système	e basé sur l'adaptation de la gamme dynamique	58
	3.4.2	Système	HDR basé sur plusieurs capteurs d'image	60
3.5	Artéfa	icts géné	rés par la suppression de ghost	61
	3.5.1	Métrique	es pour la suppression de ghost	62
		3.5.1.1	Métrique de fusion	62
	<ul> <li>3.1</li> <li>3.2</li> <li>3.3</li> <li>3.3</li> <li>3.4</li> <li>3.5</li> </ul>	<ul> <li>3.1 Introd 3.1.1</li> <li>3.2 Détec 3.2.1</li> <li>3.3 Détec 3.3.1</li> <li>3.3.2</li> <li>3.3.2</li> <li>3.3.3</li> <li>3.3.3</li> <li>3.3.4</li> <li>3.4.1</li> <li>3.4.2</li> <li>3.5 Artéfa 3.5.1</li> </ul>	<ul> <li>3.1 Introduction . 3.1.1 La défini</li> <li>3.2 Détection de g 3.2.1 Détection</li> <li>3.3 Détection et su 3.3.1 Méthode férence 3.3.1.1 3.3.2 Méthode 3.3.2.1 3.3.2.2 3.3.2.3 3.3.2.4 3.3.2.5 3.3.2.6 3.3.2.7 3.3.2.8 3.3.2.6 3.3.2.7 3.3.2.8 3.3.2.9 3.3.3 Méthode 3.3.3.1 3.3.2.9 3.3.3 Méthode 3.3.3.1 3.3.2.9 3.3.4 La supp 3.3.4.1 3.3.4.2 3.4.1 Systèmes 3.4.2 Systèmes</li> <li>3.5.1 Métrique 3.5.1.1</li> </ul>	<ul> <li>3.1 Introduction <ul> <li>3.1.1 La définition de ghost</li> <li>3.2 Détection de ghost</li> <li>3.2.1 Détection et suppression du mouvement de la caméra</li> <li>3.2.1 Détection et suppression du mouvement des objets dans la scène</li> <li>3.3.1 Méthodes de correction de ghost sans utilisation d'image de référence</li> <li>3.3.1.1 Détection basée sur le gradient</li> <li>3.3.2 Méthodes de correction utilisant une seule image de référence</li> <li>3.3.2.1 détection basée sur le gradient</li> <li>3.3.2.2 Détection basée sur le variance</li> <li>3.3.2.3 Détection basée sur l'entropie</li> <li>3.3.2.4 Détection basée sur le suppression de ghost sans utilisation d'image de référence</li> <li>3.3.2.5 Détection basée sur un seuillage multi-niveaux</li> <li>3.3.2.6 Détection basée sur des densités de probabilités conjointes</li> <li>3.3.2.7 Detection basée sur des densités de probabilités conjointes</li> <li>3.3.2.8 Détection de ghost par le calcul du flux optique</li> <li>3.3.2.9 Détection basée sur l'ordre du pixel</li> <li>3.3.3.1 Détection basée sur l'ordre du pixel</li> <li>3.3.3.2 Détection basée sur la prédiction</li> <li>3.3.4.1 Méthode basée sur l'estimation de mouvement par bloc</li> <li>3.3.4.1 Méthode basée sur l'estimation de mouvement par bloc</li> <li>3.3.4.1 Méthode basée sur l'estimation de mouvement par bloc</li> <li>3.4.2 Méthode basé sur l'agante d'optimisation basé sur un patch temporellement cohérent</li> </ul> </li> <li>3.4.1 Système basé sur l'adaptation de la gamme dynamique</li> <li>3.4.2 Système HDR basé sur plusieurs capteurs d'image</li> <li>3.5.1.1 Métrique de fusion</li> </ul>

3.6	Conclusion		66
~ ~	<b>A I I</b>		~~
	3.5.1.5	Métrique unifiée de la qualité de la suppression de ghost	65
	3.5.1.4	Métrique de gamme dynamique	64
	3.5.1.3	Métrique de différence visuelle	64
	3.5.1.2	Métrique de cohérence de gradient	63

## 3.1/ INTRODUCTION

La création d'images de grande gamme dynamique (HDR) utilisant la technique de multiples expositions issues d'un seul capteur d'image LDR est actuellement bien maîtrisée. Cependant, en présence de mouvements de la caméra ou des objets dans la scène filmée, cette génération peut devenir très difficile. Le défi se situe dans la création d'une image HDR sans que les artefacts liés au mouvement de la caméra ou ceux des objets dans la scène n'apparaissent comme des fantômes (Ghost) dans l'image HDR. Afin de remédier à cette limitation, plusieurs méthodes sont proposées dans l'état de l'art, souvent appelées méthodes ou algorithmes de deghosting. Chacune de ces méthodes aborde l'artefact de ghost à partir de différentes perspectives. Ce qui fait que les solutions proposées par chacune de ces méthodes diffèrent au niveau de la complexité algorithmique. Les différents algorithmes, que l'on peut trouver, proposent des solutions pour, d'une part, bien localiser les régions qui peuvent être affectées par l'artefact de ghost et, d'autre part, corriger cet artefact après l'avoir localisé. Les méthodes de deghosting se divisent en trois grandes familles selon la manière dont elles corrigent le ghost. La première famille tend à supprimer tous les objets en mouvement dans la scène, ce qui fait que l'image HDR résultante ne contient que la partie statique de la scène filmée (en analysant les images LDR elle discrimine les objets en mouvements, ce qui revient à créer une image HDR de l'arrière plan). La deuxième famille essaie de garder une seule occurrence de chacun des objets en mouvement dans la scène, l'objet est inséré à partir de l'image LDR la mieux exposée ce qui fait que la gamme dynamique des régions impactées par le ghost demeure faible. La troisième est la famille la plus complète mais la plus complexe, elle crée une image HDR spatialement uniforme en gamme dynamique, elle cherche à augmenter la gamme dynamique de toute l'image y compris les régions impactées par le ghost. L'augmentation de la gamme dynamique des régions qui contiennent des mouvements se fait par recalage des images et des objets en mouvement.

Dans ce chapitre on détaillera les différents algorithmes permettant la détection des régions contenant le mouvement, et les trois familles de correction de ghost. On montrera les points forts et les limitations des différents algorithmes de suppression de ghost tout en donnant les métriques permettant une comparaison quantitative et qualitative de ces différentes méthodes.

## 3.1.1/ LA DÉFINITION DE GHOST

Le terme ghost provient de l'anglais, il signifie fantôme. C'est l'apparition d'un objet plusieurs fois à des localités différentes dans l'image HDR résultante. Ces localités correspondent aux différentes positions de l'objet ou des objets en mouvement dans les images LDR qui sont utilisées dans la génération de l'image HDR correspondante. L'objet ou les objets en mouvement apparaissent intégrés ou fondus dans l'arrière plan de l'image avec une transparence comme le montre la figure 3.1. L'artefact de ghost est donné dans les cadres blancs.



FIGURE 3.1 – Apparition de ghosts suite au mouvement des objets dans les images LDRs [An et al., 2014]

# 3.2/ DÉTECTION DE GHOST

La détection de l'artefact de ghost est basée principalement sur la détection du mouvement dans les images LDRs utilisées dans la reconstruction d'une image HDR. Bien que la détection du mouvement soit une problématique classique dans le domaine de l'imagerie conventionnelle que l'on peut résoudre avec une simple soustraction entre les différentes images, elle s'avère compliquée lorsque il s'agit de l'appliquer dans le domaine de l'imagerie à grande gamme dynamique. En effet, une simple soustraction entre les images LDRs donne une valeur combinée de la variation de l'exposition et du mouvement. C'est pourquoi les méthodes de détection du mouvement doivent prendre en compte le fait que les différentes images LDRs sont acquises en utilisant des temps d'expositions différents. On trouve ainsi plusieurs méthodes de détection du mouvement appliquées dans le domaine de l'imagerie HDR. Elles sont divisées en deux catégories : la première détecte le mouvement de la caméra (mouvement global) et la seconde le mouvement des objets dans la scène (mouvements locaux).

## 3.2.1/ DÉTECTION ET SUPPRESSION DU MOUVEMENT DE LA CAMÉRA

L'existence de décalage (désalignement) entres les images LDRs qui participent à la création d'une image HDR introduit un flou global sur l'image résultante, illustré sur la figure 3.2. Pour remédier à cette limitation, la caméra d'acquisition doit être placée sur un trépied ou bien les algorithmes de création HDR doivent être précédés par une étape de détection du mouvement de la caméra afin d'aligner les différentes images LDRs avant la reconstruction HDR. Le premier algorithme de recalage d'images dans le domaine de l'imagerie HDR a été introduit par Mann et al. [Mann et al., 2002] en 2002. L'algorithme proposé utilise les équations comparamétriques (g = f(2q) = 1.68f : où q est la radiance et 2 est le ratio comparamétrique) pour aligner les images dans le contexte d'acquisition de la vidéo panoramique et l'amélioration de la dynamique n'est présente que dans les parties d'images qui se chevauchaient. La méthode la plus citée et utilisée qu'on trouve dans l'état de l'art est proposée par Ward [Ward, 2003]. Son algorithme est basé sur le calcul d'une carte binaire obtenue en faisant un seuillage par rapport à la médiane (MTB : acronyme anglais de Median Threshold Bitmap). Le MTB  $M_i$  d'une image Li est donné par l'équation 3.1.

$$M_i(p) = \begin{cases} 1 & \text{Si } L_i(p) < u_i \\ 0 & \text{Autrement} \end{cases} \text{[Ward, 2003]}$$
(3.1)

Où  $u_i$  est la valeur médiane des intensités dans l'image  $L_i$ . Les images  $L_i$  peuvent êtres approximées aux canaux vert des images LDR ou bien les images à niveaux de gris équivalentes calculées en utilisant l'équation 3.2.



## $Gris = (54 \times rouge + 183 \times vert + 19 \times bleu) [Ward, 2003]$ (3.2)

FIGURE 3.2 – Apparition de ghost suite au mouvement de la caméra.

L'algorithme permet une correction rapide du mouvement en translation de la caméra. La correction est réalisée en alignant les différentes cartes binaires obtenues à partir des images LDRs (des différentes expositions) en entrée par rapport à une image de référence sélectionnée au préalable. La correction commence par la détection du décalage entre les différentes cartes des images LDRs par rapport à la carte de l'image de référence en utilisant une image pyramidale (multi-échelle). Une autre méthode est proposée par T.Grosch [Grosch, 2006] utilisant le même principe que celle de Ward mais en incluant les rotations dans les images LDRs. Il existe aussi d'autres méthodes de recalage d'images basées sur l'extraction et la correspondance de points clés (détection des coins d'Harris [Harris et al., 1988] ou les points caractéristiques SIFT [Lowe, 2004]). La détection de mouvement de la caméra et sa correction n'est pas donnée en détail, car notre système de vision HDR (caméra HDR) sera fixe.

# 3.3/ DÉTECTION ET SUPPRESSION DU MOUVEMENT DES OBJETS DANS LA SCÈNE

Il existe plusieurs méthodes qui permettent de détecter les mouvements des objets dans les images LDR qui contribuent dans la génération d'une seule image HDR. La majorité des méthodes apparues avant 2015 sont citées dans les deux états de l'art [Srikantha et al., 2012a] [Tursun et al., 2015] dédiés à la suppression de ghost en imagerie HDR. Ces méthodes doivent traiter les images LDR en entrée en prenant en compte la variation de l'exposition d'une image à l'autre. Ces méthodes peuvent être classées selon différentes critères (domaine de fusion, carte de ghost, réglages des paramètre et seuils, etc). Dans ce manuscrit on a choisi de les classer en 3 classes selon le nombre d'images de références exigées dans la génération HDR des zones de mouvement (les images de référence dans ce classement sont celles utilisées lors de la correction de l'artefact et non plus celles utilisées afin de déterminer la localité du mouvement comme c'est le cas du classement donné dans l'état de l'art de Srikantha [Srikantha et al., 2012a]). La première classe ne nécessite aucune image de référence, la deuxième exige une image et cette dernière doit être la mieux exposée parmi les images LDR, et la dernière classe prend un nombre variable d'images selon les objets en mouvement.

# 3.3.1/ MÉTHODES DE CORRECTION DE GHOST SANS UTILISATION D'IMAGE DE RÉFÉRENCE

Ce sont des méthodes qui suppriment l'objet en mouvement de l'image HDR finale. Dans le cas où ces méthodes gardent une occurrence de l'objet en mouvement, elles doivent impérativement sélectionner une image de référence à partir de laquelle elles récupèrent cette occurrence.

#### 3.3.1.1/ DÉTECTION BASÉE SUR LE GRADIENT

L'hypothèse de cette méthode [Zhang et al., 2010] est que la direction du gradient change dans les régions qui contiennent du mouvement. Elle utilise une mesure de cohérence basée sur le changement de la direction du gradient entre les différentes expositions et l'image de référence. Cette mesure de cohérence est utilisée ensuite comme une fonction de pondération dans l'équation de la génération d'images HDR. Pour chaque image LDR en entrée, l'information du gradient est extraite par une convolution avec la première dérivée d'un noyau Gaussien 2*D*. Les différentes étapes de cette méthode sont données maintenant.

1. Calcul de la visibilité d'un pixel

La visibilité relative d'un pixel (i, j) dans une exposition  $L_k$  est donnée par l'équation 3.3.

$$V_{ij}^{k} = \frac{M_{ij}^{k}}{\sum_{l=1}^{N} M_{ij}^{l} + \epsilon}$$
(3.3)

Où  $M_{ij}^k$  est l'amplitude du gradient et  $\epsilon$  est une valeur faible ajoutée dans le but d'éviter la division par zéro (la singularité).

 La mesure de la cohérence : C'est un score défini par l'accumulation des changements de direction de gradient de chaque exposition pour refléter sa cohérence dans toute la séquence d'images LDR à fusionner.

$$S_{ij}^{k} = \sum_{l=1}^{N} \exp\left(\frac{-(d_{ij}^{kl})^{2}}{2\sigma_{s}^{2}}\right)$$
(3.4)

 $d_{ij}^{kl}$  est la direction de gradient calculée entre  $L_k$  et  $L_l$  à la position (i, j) et  $\sigma$  est l'écart type sur une fenêtre de taille  $(2r+1) \times (2r+1)$  (avec  $r = \{1, 2, ...\}$  une variable qui fixe la taille de masque).

3. La direction du gradient

$$d_{ij}^{kl} = \frac{\sum_{x=-r}^{r} \left| D_{(i+x)(j+x)}^{k} - D_{(i+x)(j+x)}^{l} \right|}{(2r+1)^{2}}$$
(3.5)

Où  $D_{(i+x)(j+x)}^k$  est la direction du gradient calculée dans une image  $L_k$  pour une fenêtre de taille  $(2r+1)^2$ 

4. Le score ré-affiné

C'est une valeur introduite dans le but d'éliminer l'effet des régions saturées. Le score ré-affiné est calculé par la fonction 3.6.

$$C_{ij}^{k} = \frac{S_{ij}^{k} \times \omega_{ij}^{k}}{\sum_{l=1}^{N} (S_{ij}^{l} \times \omega_{ij}^{l}) + \epsilon}$$
(3.6)

Où  $S_{ij}^{l}$  est la la mesure de la cohérence donnée par l'équation 3.4, et  $\omega_{ij}^{l}$  est un poids de pondération des mesures de la cohérence calculée en utilisant la fonction donnée en équation 3.7 et  $\epsilon$  est une valeur faible utilisée pour éviter la singularité.

$$\omega_{ij}^{k} = \begin{cases} 1 & \text{Si } 25 \le Z_{ij}^{k} \le 255 \\ 0 & \text{Autrement} \end{cases}$$
(3.7)

## 5. La fonction de pondération HDR

Les poids de pondération utilisés lors de la génération HDR permettant une suppression efficace de l'artefact de ghost sont donnés par la fonction 3.8. Ils sont calculés en utilisant la visibilité  $V_{ij}^k$  et le score ré-affiné  $C_{ij}^k$ .

$$W_{ij}^{k} = \frac{V_{ij}^{k} \times C_{ij}^{k}}{\sum_{l=1}^{N} (V_{ij}^{l} \times C_{ij}^{l}) + \epsilon}$$

$$(3.8)$$

#### 6. La génération HDR

La création de l'image HDR finale se fait en utilisant l'équation 3.9

$$I_{ij} = \sum_{K=1}^{N} W_{ij}^{k} \times Z_{ij}^{k}$$
(3.9)

## 3.3.2/ Méthodes de correction utilisant une seule image de référence

L'image de référence est utilisée pour détecter le mouvement et aussi pendant la correction. Ces méthodes choisissent généralement l'image qui représente le nombre minimal de pixels saturés et sous exposés comme image de référence car c'est l'image qui contient le plus d'information parmi la liste d'images LDR.

## 3.3.2.1/ DÉTECTION BASÉE SUR LA VARIANCE

La détection des régions impactées par l'artefact de ghost se fait dans le domaine de la radiance, en calculant une image de la variance 3.10.

$$VI_{ij} = \frac{\sum_{k=1}^{n} w(Z_{ij}^{k}) (E_{ij}^{k})^{2} / \sum_{k=1}^{n} w(Z_{ij}^{k})}{(\sum_{k=1}^{n} w(Z_{ij}^{k}) E_{ij}^{k})^{2} / (\sum_{k=1}^{n} w(Z_{ij}^{k}))^{2}} - 1$$
(3.10)

L'image de la variance est obtenue en évaluant la variance de l'irradiance au niveau pixel [Jacobs et al., 2008] des différentes images LDRs. Les régions impactées par le mouvement sont détectées par leurs variances élevées par rapport à un seuil déterminé au préalable comme le montre la figure 3.3. La carte de ghost est calculée par comparaison des variances de chaque pixel à un seuil global 3.11. Pour les images couleurs,



FIGURE 3.3 – L'image de variance calculée pour trois images à différentes expositions. [Jacobs et al., 2008].

l'image de la variance VI est calculée en utilisant le maximum des trois canaux couleurs.

$$G_{ij} = \begin{cases} 1 \text{ Si } VI_{ij} \ge S \text{ euil} \\ 0 \text{ Autrement} \end{cases}$$
(3.11)

Le seuil est fixé à 0.18 dans [Jacobs et al., 2008] pour une image de variance VI normalisée.

La génération de l'image HDR finale se fait en utilisant l'équation de fusion HDR générale pour les pixels non impactés par le ghost (tous les pixels des différentes images LDRs contribuent dans cette génération). Pour les zones touchées par le mouvement, cette méthode propose l'utilisation des pixels de l'image LDR représentant le minimum de zones saturées seulement, tout en favorisant une image ayant une longue exposition lorsque ceci est possible. Ce choix se justifie par le fait que cette dernière présente le meilleur SNR en sein d'une faible saturation et d'un faible bruit.

## 3.3.2.2/ DÉTECTION BASÉE SUR L'ENTROPIE

Cette méthode a été introduite par Jacobs et al. [Jacobs et al., 2008] comme une amélioration de leur méthode basée sur la variance qui est limitée lorsque les objets en mouvement ont la même couleur que l'arrière plan donc un faible contraste. Le calcul de l'entropie permet selon Jacobs et al. [Jacobs et al., 2008] une détection précise des mouvements dans les régions de faible contraste. La méthode commence par le calcul d'un voisinage local basé sur une carte d'entropie calculée pour chaque image LDR. L'entropie 3.12 est calculée à partir d'un histogramme local pour chaque pixel (i, j) dans l'image  $L_k$ . L'histogramme local est calculé dans une fenêtre d'une taille de  $(2r+1) \times (2r+1)$  (avec  $r = \{1, 2, ...\}$  une variable qui fixe la taille de masque ou de la fenêtre) entourant le pixel (i, j).

$$H_{ij}^{k} = -\sum_{x=0}^{B-1} P(X = x) \log(P(X = x))$$
(3.12)

Le nombre *B* représente le nombre de barres de l'histogramme. La probabilité (X = x) est calculée à partir de l'histogramme normalisé. Lorsque cette probabilité est nulle, l'équation 3.12 est nulle aussi pour éviter un résultat complexe du logarithme.

Après avoir calculé l'entropie pour chaque image LDR, une image d'incertitude est calculée à partir de la différence pondérée des entropies des différentes images LDR en utilisant l'équation 3.13. Pour les images couleurs, l'image de l'incertitude *UI* est calculée en utilisant le maximum des trois canaux couleurs.

$$UI_{ij} = \sum_{k=1}^{N} \sum_{l=1}^{l < k} \frac{\nu^{kl}}{\sum_{k=1}^{N} \sum_{l=1}^{l < k} \nu^{kl}} h_{ij}^{kl}$$
(3.13)

Où la différence d'entropie  $h_{ij}^{kl}$  est donnée par  $|H_{ij}^k - H_{ij}^l|$ ,  $v^{kl} = min(w(Z_{ij}^k), w(Z_{ij}^l))$  et  $w(Z_{ij}^k)$  est le poids donné par la fonction de pondération 3.14.

$$w(Z_{ij}^k) = \begin{cases} (Z_{ij}^k \times 0.9/127) + 0.05 & \text{Si } Z_{ij}^k \le 127 \\ (255 - Z_{ij}^k) \times 0.9/127) + 0.05 & \text{Si } Z_{ij}^k > 127 \end{cases}$$
(3.14)

La carte de ghost est calculée en faisant un seuillage et une binarisation de l'image d'incertitude 3.15.

$$G_{ij} = \begin{cases} 1 \text{ Si } UI_{ij} \ge S \text{ euil} \\ 0 \text{ Autrement} \end{cases}$$
(3.15)

Le seuil est fixé à 0.7 dans l'exemple donnée par Jacobs et al. [Jacobs et al., 2008], l'image d'incertitude UI est normalisée et calculée à partir d'images d'entropie obtenues en utilisant un rayon r = 40 et B = 200.

L'image HDR finale est crée de la même manière que pour la méthode de détection par la variance en utilisant une seule image LDR dans la génération HDR pour les pixel impactés par le mouvement.



FIGURE 3.4 – Les cartes multi-seuils des deux images faiblement exposées (haut à gauche) et hautement exposées (haut à droite) en prenant l'image de milieu comme référence [Min et al., 2009].

## 3.3.2.3/ DÉTECTION BASÉE SUR UN SEUILLAGE MULTI-NIVEAUX

Elle détecte les zones représentant des mouvements dans la scène en se basant sur une carte de seuillage multi-niveaux [Min et al., 2009]. La condition imposée par cette méthode est que les niveaux du gris à des localités particulières doivent présenter une propriété croissante quand les images LDR sont acquises de la basse à la plus haute exposition. Chacune des images LDR  $I_j$  est divisée en N parties, où chaque partie est seuillée à un niveau (le nombre de niveaux est égal au nombre de parties), toutes les parties contiennent le même nombre de pixels. La carte des seuils multi-niveaux  $L_{i,j}$  (j = 1, ...N) est calculée par classification des valeurs d'intensité de  $I_j$  en N niveaux en utilisant ces seuils. L'estimation de la carte de ghost est générée en utilisant la carte multi-seuils (voir la figure 3.4 ) selon la fonction 3.16.

$$M_{i,j} = \begin{cases} 1 \text{ Si } |L_{i,ref} - L_{i,j}| \ge 1, k \neq ref \\ 0 \text{ Autrement} \end{cases}$$
(3.16)

Où  $L_{i,j}$  est la carte des seuils multi-niveaux, ref est l'index de l'image de référence choisie comme l'image d'exposition moyenne. Dans l'exemple donné dans [Min et al., 2009], une série de 3 images LDR est utilisée avec le choix de l'image numéro 3 comme image de référence.

Pour ré-affiner la détection des objets en mouvement, cette méthode utilise la carte d'er-

reurs binaires  $Er_{i,j}$  (voir l'équation3.17) qui repère les fausses détections qui peuvent être introduites lorsque les intensités des pixels sont similaires ou proches de la valeur de seuil.

$$Er_{i,j} = \begin{cases} 1 \text{ pour } T_{j,k} - 1 \le I_{i,j} \le T_{j,k} + 1, 1 \le k \le N \\ 0 \text{ Autrement} \end{cases}$$
(3.17)

Où  $I_{i,j}$  est l'intensité au pixel *i* de l'image LDR *j* (*j* est l'ordre d'exposition),  $T_{j,k}$  est le  $k^{eme}$  seuil utilisé pour l'extraction de la carte de seuil multi-niveaux.

Cette méthode supprime l'artefact de ghost en ajoutant un facteur  $ME_{i,j}$  (équation 3.19) qui sélectionne les pixels des images LDR qui contribuent dans l'image de radiance finale lors de la génération HDR.

$$\ln E_{i} = \frac{\sum_{j=1}^{P} \omega(Z_{i,j}) (ME_{i,j}) \left( g(Z_{i,j}) - \ln \Delta t_{j} \right)}{\sum_{j=1}^{P} \omega(Z_{i,j}) (ME_{i,j})}.$$
(3.18)

Où  $E_i$  est la radiance du pixel *i* dans l'image (la matrice de l'image est transformée en un vecteur unidimensionnel d'indice *i*), *P* est le nombre d'images LDR,  $Z_i$ , *j* est la valeur (R, V et B) du pixel *i* dans l'image *j*, *g*() est le logarithme de la fonction inverse de la CRF,  $\Delta t_j$  est le temps d'exposition de l'image LDR *j* et  $\omega$ () est le poids de pondération de Débévec.

$$ME_{i,j} = \begin{cases} 1 \text{ pour } j = ref \\ 1 \text{ pour } M_{i,j} = 0, j \neq ref \\ 0 \text{ pour } M_{i,j} = 1, j \neq ref \\ 0.2 \text{ pour } Er_{i,j} | Er_{i,ref} = 1, j \neq ref \end{cases}$$
(3.19)

## 3.3.2.4/ DÉTECTION BASÉE SUR UN BITMAP

Cette méthode [Pece et al., 2010] se base sur la relation d'un pixel avec la médiane des intensités de l'image et ceci pour chaque exposition. Cette relation doit être la même pour toutes les images LDR. Le bitmap binaire médiane  $M_k$  est obtenu par un seuillage de l'image LDR  $L_k$  basé sur sa valeur de pixel médiane. Les régions de  $M_k$  sombre ( $M_k = 0$ ) indiquent les pixels dont la valeur est inférieure à la médiane et  $M_k = 1$  pour les pixels dont la valeur est supérieure à la valeur médiane comme montré dans l'image 3.5.

Pour détecter le mouvement, la méthode basée sur le MTB fait la somme (au niveau pixel) de tous les seuillages réalisés par rapport à la médiane  $M_{ij}^k$  3.20. pour les pixels



FIGURE 3.5 – Le calcul de bitmap de 2 images statiques à différentes expositions donne un même résultat[Pece et al., 2010].

des régions qui ne représentent pas du mouvement. Cette somme est égale à 0 où à N (nombre d'images LDR en entrée) car la relation des pixels n'a pas changé par rapport à la médiane. Si la somme est différente de 0 et de N, le pixel concerné est considéré comme pixel d'une région de ghost 3.21.

$$S_{ij} = \sum_{k=1}^{N} M_{ij}^{k}$$
(3.20)

$$G_{ij} = \begin{cases} 0 & \text{Si } S_{ij} = 0 \text{ ou } S_{ij} = N \\ 1 & \text{Autrement} \end{cases}$$
(3.21)

Cette méthode a été proposée pour la génération HDR dans le domaine de l'image. Elle utilise la méthode proposée par Martens [Mertens et al., 2009]. Pour supprimer l'artefact de ghost, la méthode basée sur la MTB n'utilise que les pixels de l'image bien exposée pour les régions présentant le mouvement.

## 3.3.2.5/ DÉTECTION BASÉE SUR RANSAC

Pour deux images de différents temps d'exposition, la méthode proposée par Gallo [Gallo et al., 2009], suppose que les valeurs d'intensités à chaque localité (i, j) de deux images sont liées par la relation donnée dans l'équation 3.22. Pour les zones non saturées des deux images, cette relation change seulement si les zones sont touchées par le mouvement.

$$ln(Z_{ij}^k) - ln(\Delta t_k) = ln(Z_{ij}^l) - ln(\Delta t_l)$$
(3.22)

Pour une robustesse au bruit, le traitement est réalisé par partie.

- Au début, les régions saturées des différentes images LDR sont calculées et l'image LDR qui représente le nombre minimum de régions saturées est élue image de référence.
- Pour déterminer si une région de taille (r × r) dans une exposition L<sub>k</sub> contient des mouvements, la méthode demande de tracer le log des intensités de la région dans L<sub>k</sub> et de les comparer à celui calculé pour la région correspondante dans l'image de référence.

La génération de l'image HDR est réalisée en utilisant la méthode de Débévec, et ceci en utilisant tous les patchs des différentes images LDR qui ne sont pas impactés par le mouvement. Pour les région touchées par le mouvement, seuls les patchs consistants avec celui de l'image de référence seront utilisés dans la génération HDR.

#### 3.3.2.6/ DÉTECTION BASÉE SUR DES DENSITÉS DE PROBABILITÉS CONJOINTES

Cette méthode propose une approche basée sur des probabilités (par le calcul des histogrammes conjoints 3.23 ) afin de détecter le mouvement dans la scène [Heo et al., 2010]. Cette méthode commence par sélectionner une image de référence qui sera utilisée dans l'estimation de la relation entre les différentes expositions et cette dernière. L'estimation est réalisée par la construction d'un histogramme conjoint  $P_{n_0,n}^k$  pour chaque canal de couleur  $k \in (R, VetB)$  entre l'image de référence  $n_0$  et les autres images LDR n. Ces histogrammes sont nommés fonctions de densité de probabilité conjoints (PDFS).

$$P_{n_0,n}^k(i,j) = \frac{1}{M} \sum_{P} G_n(p) \times T\left[(i,j) = (I_{n_0}^k(p), I_n^k(p))\right]$$
(3.23)

Où *p* est l'index de position d'un pixel, T[.] = 1 si l'argument est juste et 0 autrement, *M* est le nombre total des pixels correspondants dans une image,  $I_{n_0}^k(.)$  est la valeur d'une intensité du canal *k* de l'image *n* et  $G_n(.)$  est la fonction de poids de ghost. A la première itération  $G_n(.)$  est mis à 1 pour tous les pixels de toutes les expositions puis une estimation par noyau (dite aussi fenêtrage de "Parzen" [Parzen, 1962]) est effectuée par convolution avec une gaussienne 2*D* pour avoir des fonctions de densité de probabilité conjointes lisses. L'estimation de la région de ghost se fait en utilisant la relation  $G_n(.)$  définie dans l'équation 3.24 pour toutes les images LDR sauf l'image LDR de référence.

$$G_{n}(p) = \begin{cases} 0 & \text{Si } P_{n_{0},n}^{R}(I_{n_{0}}^{R}(p), I_{n}^{R}(p)) < C \\ & ou \\ P_{n_{0},n}^{V}(I_{n_{0}}^{V}(p), I_{n}^{V}(p)) < C \\ & ou \\ P_{n_{0},n}^{B}(I_{n_{0}}^{B}(p), I_{n}^{B}(p)) < C \\ 1 & \text{Autrement} \end{cases}$$
(3.24)

Où *C* est un seuil fixé à  $10^{-5}$ . Pour ré-affiner le calcul de détection de ghost, cette méthode utilise une approche de minimisation d'énergie pour chaque image. Elle définit l'énergie totale à minimiser en utilisant l'équation 3.25.

$$E(f_n) = \sum_p D_p(f_n(p)) + \sum_p \sum_{q \in N(p)} V_{qp}(f_n(p), f_n(q))$$
(3.25)

Où  $f_n \in \{0, 1\}$  est un booléen qui prend les valeurs 0, 1 quand le pixel est un ghost ou non respectivement dans l'image n, N(p) sont les pixels de voisinage du pixel p et  $D_p()$  est le coût de données défini dans la relation 3.26.  $V_{qp}()$  est le coût de lissage défini comme une fonction linéaire tronquée équation 3.27.

$$D_{p}(f_{n}(p)) = \begin{cases} 0 & \text{Si } f_{n}(p) = 0 \land G_{n}(p) = 0 \\ & ou \\ f_{n}(p) = 1 \land G_{n}(p) = 1 \\ \beta & \text{Autrement} \end{cases}$$
(3.26)

Où  $\beta$  est une constante fixée à 2.5 dans l'exemple fourni dans [Heo et al., 2010].

$$\lambda_{pq} = \begin{cases} \lambda_L & \text{Si } |I_{n_0}(p) - I_{n_0}(q)| < \eta \lor |I_n(p) - I_n(q)| < \nu \\ \lambda_S & \text{Autrement} \end{cases}$$
(3.27)

Où  $\lambda_L > \lambda_S$  et  $I_{n_0}(p)$  représente la valeur en niveau du gris de pixel p dans l'image n

La suppression de ghost dans l'image HDR finale est réalisée en modifiant la fonction de pondération de la génération HDR par l'ajout du terme de poids de ghost  $G_n(p)$  dans l'équation 3.28.

$$ln\overline{E}_{n}^{k}(p) = \frac{1}{C} \sum_{p, I_{n}^{k}(p) = x} G_{n}(p) ln E^{k}(p)$$
(3.28)

Ou *C* est une constante de normalisation pour avoir une courbe plus lisse,  $E_n^k$  est l'irradiance d'un pixel *p* dans l'image *n* 

La méthode est améliorée par l'utilisation d'une fonction de transfert d'intensité globale obtenue à partir des fonctions de densité de probabilité conjointe entre les différentes images LDR. Pour estimer une valeur d'irradiance fiable, elle utilise une fonction de filtrage pondérée généralisée utilisant la fonction de transfert d'intensité globale. Cette dernière est calculée entre l'image de référence  $n_0$  et l'image n en utilisant la méthode des moindres carrés (MMSE) et les PDFS conjointes estimées (plus de détails dans [Heo et al., 2010]).

#### 3.3.2.7/ DETECTION BASÉE SUR LA COMPENSATION DU MOUVEMENT

Cette méthode [Menzel et al., 2007] est basée sur la connaissance du mouvement entre les différentes images LDR afin de détecter les régions impactées par l'artefact de ghost. Elle commence par un alignement non linéaire des images LDR pour lequel elle choisit d'abord une image de référence r en sélectionnant celle qui contient l'entropie la plus élevée et ensuite toutes les autres images sont alignées sur celle-ci en utilisant des macro-blocs avec une maximisation de la corrélation croisée. Pour capturer les différentes images LDR sans que l'appareil photo soit fixé sur un trépieds. Cette méthode propose d'utiliser la fonction de bracketing d'exposition automatique (AEB) de l'appareil photo. La fonction AEB permet à l'appareil photo de capturer une séquence d'images à des expositions optimales (sous-exposées et surexposées). Cette méthode se limite à trois expositions seulement car au delà de trois images, il est difficile de trouver une image de référence optimale selon l'algorithme proposé. Cette limitation est liée à l'espacement temporel considérable qui contraint la corrélation optimale car celle-ci ne peut être attendue que pour les images voisines. L'idée de base de l'estimation du mouvement proposé est que pour un macro-bloc M donné dans l'image de référence, le macro-bloc déplacé de  $\delta$  correspondant dans l'image *i* aura la cross-corrélation maximale localement  $C_i(M,\delta)$  3.29.

$$C_{i}(M,\delta) = \frac{\sum_{p \in M} \tilde{c}_{i}(p+\delta).\tilde{c}_{r}(p)}{N(M)\sqrt{\sum_{p \in M} \left\|\tilde{c}_{i}(p+\delta)\right\|^{2}}}$$
(3.29)

Avec 
$$\tilde{c}_i(p+\delta) = c_i(p+\delta) - \begin{pmatrix} 1\\1\\1 \end{pmatrix} \frac{\sum_{p \in M} ||c_i(p+\delta)||_1}{3N(M)}$$
 et  $\tilde{c}_r(p) = c_r(p) - \begin{pmatrix} 1\\1\\1 \end{pmatrix} \frac{\sum_{p \in M} ||c_r(p)||_1}{3N(M)}$ 

Où  $c_k(p)$  est le vecteur de couleur RVB de l'image LDR k au pixel p, et N(M) est le nombre de pixels contenus dans le macro-bloc.

Pour supprimer le ghost, cette méthode utilise un indice de confiance  $K_i(p)$  (voir l'équation 3.30) pour discriminer les blocs de pixels susceptibles de contenir l'artefact de ghost. Pour chaque pixel de toutes les images LDRs (à l'exception des pixels de l'image de référence r), une valeur de confiance dérivée de la cross-corrélation est calculée. Un petit voisinage  $N_p$  autour du pixel p est extrait de chaque image i et l'indice de confiance  $K_i(p)$ est calculée à partir de ces images. Cet indice est utilisé comme une condition dans l'équation 3.31 dans la fonction de pondération traditionnelle de Debevec.

$$k_{i}(p) = \frac{\sum_{j \in N_{p}} c_{i}(j).c_{r}(j)}{\sqrt{\sum_{j \in N_{p}} ||c_{i}(j)||^{2}} \sqrt{\sum_{j \in N_{p}} ||c_{r}(j)||^{2}}}$$
(3.30)

$$\omega_i^*(p) = \begin{cases} 0 & \text{Si } k_i(p) < t_k \text{ avec } t_b < c_i(p) < t_s \\ & ou \\ \omega(p) & \text{Autrement} \end{cases}$$
(3.31)

Dans l'exemple donné par l'auteur de cette méthode, le voisinage utilisé est de  $11 \times 11$  pixels centrés sur le pixel courant et les valeurs de seuil  $t_K$  et  $t_s$  sont fixés à 0,95 et 0,05 respectivement.

## 3.3.2.8/ DÉTECTION DE GHOST PAR LE CALCUL DU FLUX OPTIQUE

L'algorithme proposé par Kang et al. [Kang et al., 2003] estime le mouvement en deux étapes. Tout d'abord, les auteurs alignent les images LDR  $S^-$  et  $S^+$  en estimant une transformée affine qui aligne l'une sur l'autre. Ensuite, ils utilisent un flux optique basé sur le gradient pour calculer un champ de mouvement dense qui apporte des corrections locales à la transformation globale estimée au paravent. Cette méthode consiste aussi à calculer un champ d'écoulement bidirectionnel à partir de l'image LDR en cours (l'image LDR capturée avec le temps d'exposition médian) au lieu de calculer pour les deux autres images LDRs séparément. Ceci permet d'éviter les problèmes de remplissage de trous lors de la génération de chaque image LDR intermédiaire par interpolation. Pour chaque pixel de cette image LDR interpolée en sortie  $\tilde{S}_B$ , cette méthode obtient un vecteur composite qui pointe dans les images LDRs précédentes et suivantes. Ces vecteurs sont la somme des composantes affines et locales. La composante affine est dérivée des paramètres de distorsion globale alors que la composante locale est générée par l'algorithme de flux optique symétrique proposé par l'auteur de cette méthode [Kang et al., 2003]. Pour l'estimation du mouvement local, cette méthode s'appuie sur une variante de la technique proposée par Lucas et Kanade [Lucas et al., 1981] dans le framework laplacien pyramidal [Bergen et al., 1992] afin de gérer les cas de flux dégénérés, et ceci par le gauchissement des deux images source vers la trame intermédiaire et l'estimation des vecteurs de flux résiduels entre ces deux images déformées. L'accumulation des résidus dans la pyramide produit un champ de flux (d'écoulement) symétrique centré sur la trame intermédiaire. Cette technique est améliorée aussi par l'inclusion de flux affine global pendant le gauchissement de sorte que les résidus accumulés restent toujours représentés en termes d'une correction locale au flux global asymétrique.

La suppression de ghost dans l'image HDR finale, se fait par la modulation de la fonction de pondération dans l'équation de fusion HDR. Cette modulation est réalisée par la combinaison d'un poids donnée par la fonction 3.32 avec le poids donné par la fonction de pondération de la fusion HDR (l'auteur a utilisé la fonction de pondération de [Mitsunaga et al., 1999] nommée  $f_W$ ) selon la fonction 3.33.

$$F_M(p,q) = \begin{cases} 2(\frac{|p-q|}{\delta_{max}})^3 - 3(\frac{|p-q|}{\delta_{max}})^2 + 1 & \text{si} \quad |p-q| < \delta_{max} \\ 0 & \text{Autrement} \end{cases}$$
(3.32)

Où les pixels *p* et *q* de la même localité provenant des deux différentes images LDR en radiance,  $\delta_{max}$ est un paramètre à fixer par l'utilisateur. La génération HDR se fait en utilisant la fonction donnée par l'équation 3.33.

$$R = \frac{f_{WM}(p, p^{-})p^{-} + f_{WM}(p, p^{+})p^{+} + f_{W}(p)p}{f_{WM}(p, p^{-}) + f_{WM}(p, p^{+}) + f_{W}(p)}$$
(3.33)

Les pixels p,  $p^-$  et  $p^+$  proviennent des images LDRs de radiance actuelle, précédente et suivante  $\hat{L}$ ,  $\hat{S}_U^-$  et  $\hat{S}_U^+$ , respectivement. La fonction de pondération  $f_{WM}(p,q) = f_M(|p-q|)f_W(p_w)$  est la fonction de pondération basée sur l'intensité  $f_W$  de Mitsunaga et Nayar [Mitsunaga et al., 1999] modulée par la carte de vraisemblance  $f_M$ .

## 3.3.2.9/ DÉTECTION ET SUPPRESSION DE GHOST PAR UNE FONCTION DE PONDÉRA-TION

Cette méthode [Srikantha et al., 2012b] détecte et supprime l'artefact de ghost par l'utilisation d'une fonction de pondération simple donnée dans l'équation 3.34. Cette méthode exploite l'adaptation par pondération. Elle ajuste le poids des pixels des différentes images LDR par rapport à l'écart de ces même pixels avec ceux de l'image de référence. La fonction donne un poids plus élevé aux pixels dont la valeur est proche de la valeur de référence, et un poids faible pour les pixels dont la valeur diverge considérablement de la valeur de référence.

$$w(Z_{ij}^k) = \frac{[a(Z_{ij}^{ref})]^2}{[a(Z_{ij}^{ref})]^2 + [\frac{(f^{-1}(Z_{ij}^k) - f^{-1}(Z_{ij}^{ref}))}{f^{-1}(Z_{ij}^{ref})}]^2}$$
(3.34)

Où  $Z_{ij}^k$  est le pixel à la position (i, j) de l'image  $I^k$ , a(Z) est une fonction de valeur de pixel dans l'image LDR de référence normalisée dans l'intervalle [0, 1].

$$a(x) = \begin{cases} 0.058 + 0.68(x - 0.85) & \text{si } x \ge 0.85\\ 0.04 + 0.12(1 - x) & \text{si } x < 0.85 \end{cases}$$
(3.35)

## 3.3.3/ MÉTHODES DE DÉTECTION UTILISANT PLUSIEURS IMAGES DE RÉFÉ-RENCES

Ces méthodes utilisent différentes images LDRs comme image de référence. Ces dernières peuvent être utilisées toutes lors de la génération HDR de la région affectée par le mouvement ou une seule à la fois, et ceci selon la méthode utilisée. Dans le cas où l'approche utilisée permet le recalage et l'alignement des objet en mouvement, toutes les images de référence sont utilisées. Cependant, une seule image de référence est sélectionnée parmi la liste des images de référence pour chaque zone représentant du mouvement, le choix étant fait selon le bruit et le contraste de la région en question.

#### 3.3.3.1/ DÉTECTION BASÉE SUR L'ORDRE DU PIXEL

Elle est proposée par Sidibé [Sidibe et al., 2009]. Elle repose sur le fait que la fonction de réponse de la caméra est monotone et croissante. La monotonie de la CRF garantit

une consistance dans l'ordre des pixels. La valeur d'un pixel (i, j) pour un temps d'exposition  $\Delta t_l$  doit être inférieure à sa valeur pour une exposition  $\Delta t_k$  lorsque  $\Delta t_l < \Delta t_k$  tout en considérant que l'illumination de la scène est stable.

$$G_{ij} = \begin{cases} 0 \text{ Si } Z_{ij}^1 \le Z_{ij}^2 \le \dots \le Z_{ij}^k & \text{pour } \Delta t_1 \le \Delta t_2 \le \dots \le \Delta t_k \\ 1 \text{ Autrement} \end{cases}$$
(3.36)

Cette méthode discrimine les pixels considérés saturés ( $Z_{ij} < 20$  et  $Z_{ij} > 240$ ). Pour supprimer l'artefact de ghost, cette méthode divise la liste des images LDR en deux parties  $W_j$  et  $H_j$  de telle manière la première partie  $W_j$  contient les images LDR qui contiennent le mouvement à la position j et l'autre partie  $H_j$  les images LDR qui ne représentent pas de mouvement. Pour les régions qui contiennent le mouvement, la génération HDR ne se fait qu'avec les images de la liste  $H_j$  et pour les autres régions, elle utilise toutes les images LDR.

#### 3.3.3.2/ DÉTECTION BASÉE SUR LA PRÉDICTION

Pour détecter qu'un pixel d'une image LDR  $L_k$  appartient à une région de ghost, cette méthode (proposée par Grosch) [Grosch, 2006] se base sur la déviation de la couleur d'un pixel par rapport à la couleur prédite du même pixel à partir d'une image LDR précédente  $L_l$  en exploitant le temps d'exposition de l'image  $L_k$ . La prédiction de la valeur du pixel est réalisée par la fonction 3.37. La carte de ghost est obtenue en utilisant la fonction 3.38 qui compare la différence entre les valeurs des pixels et leurs valeurs prédites avec un seuil prédéfini au préalable.

$$\check{\mathsf{Z}}_{ij}^{l} = f\left(\frac{\Delta t_{l}}{\Delta t_{k}}f_{-1}(Z_{ij}^{k})\right)$$
(3.37)

Où f() est la fonction de réponse de la caméra,  $\Delta t_l$  est le temps d'exposition de l'image prédite  $Z_l$  et  $\Delta t_k$  est le temps d'exposition de l'image  $Z_k$  en entrée de processus de prédiction.

$$G_{ij} = \begin{cases} 1 \text{ Si } \left| \breve{Z}_{ij}^k - Z_{ij}^k \right| \ge S \text{ evil} \\ 0 \text{ Autrement} \end{cases}$$
(3.38)

L'image HDR est obtenue en fusionnant les pixels des différentes images LDR lorsqu'ils n'appartiennent pas à une région représentant du mouvement (leurs valeurs dans la carte de ghost sont à 0). Pour les régions présentant des mouvements, les pixels concernés sont choisis à partir d'une image LDR de référence directement. L'image de référence selon cette méthode correspond à l'image LDR qui présente le minimum de pixels saturés dans la région concernée par le mouvement.

#### 3.3.4/ LA SUPPRESSION DU GHOST POUR LA VIDÉO HDR

Kang et al [Kang et al., 2003]. ont proposé un framework vidéo HDR qui acquiert une série temporelle d'expositions différentes (images LDR à différentes expositions) et fusionne la radiance de la scène suivant le flux optique pour éviter les artefacts de ghost dans une vidéo HDR (voir la sous section 3.3.2.8). Plusieurs approches se sont inspirées de la méthode de Kang et al mais en utilisant des méthodes d'estimation de mouvement différentes.

#### 3.3.4.1/ MÉTHODE BASÉE SUR L'ESTIMATION DE MOUVEMENT PAR BLOC

Elle est proposée par Mangiat et Gibson [Mangiat et al., 2010], [Mangiat et al., 2011]. Cette méthode a introduit une technique d'estimation de mouvement basée sur des blocs plutôt que sur le flux optique basé sur le gradient, et ceci afin d'éviter les occlusions et détecter les mouvements rapides des objets. L'estimation du mouvement est réalisé par le calcul des vecteurs du mouvement d'avant (à partir de l'image  $I_{t-1}$ ) et d'après (à partir de l'image  $I_{t+1}$ ) l'image courante ( $I_t$ ) par le biais de logiciel de référence JM de standard H.264 et la somme des différences absolues (SAD : acronyme anglais de Sum of Absolute Differences) réalisée par blocs de taille  $16 \times 16$  pixels (des deux composante luminance et couleur). Cette estimation est améliorée dans les régions saturées en exploitant la similarité de la couleur dans les images adjacentes. Cette dernière est suivie par l'utilisation d'un filtre inter-bilatéral afin d'améliorer la détection et la suppression du mouvement et de préserver les contours et les structures dans l'image LDR d'origine et ceci par addition d'un coût supplémentaire au SAD selon l'équation 3.40. L'estimation du mouvement est réalisée entre une image LDR de faible exposition Z<sub>s</sub> et une autre de forte exposition  $Z_l$  par l'intermédiaire de l'image d'une forte exposition  $\tilde{Z}_l$  obtenue par prédiction en utilisant l'équation 3.39. Les pixels saturés (en comparaison a des seuils délimitant la plage de valeurs correcte des pixels) son considérés comme perdus. Chaque pixel perdu est recalculé à partir de son voisinage (par bloc de  $4 \times 4$ ) en utilisant la fonction donnée en équation 3.41.

$$\tilde{Z}_l = g^{-1}(g(Z_s) - \ln\Delta t_s + \Delta t_l)$$
(3.39)

Où  $\Delta t_l$ ,  $\Delta t_s$  sont les temps d'exposition long et court respectivement.

$$co\hat{u}t = SAD(I_{t-1}(x_1, y_1), I_{t+1}(x_2, y_2)) + \lambda |MV_{pred} \pm MV_k|$$
(3.40)

Où  $I_{t-1}, I_{t+1}$  sont les images LDR avant et après l'image LDR  $I_t$  respectivement,  $MV_k$  est le vecteur du mouvement,  $MV_{pred}$  est le vecteur du mouvement prédit et  $\lambda$  est une constante.

$$w(p,q_k) = r(p,q_k)c(p,q_k) = \frac{1}{\|p-q_k\|} \frac{1}{\|I_{t'}(p+MV_k) - I_{t'}(q_k+MV_k)\|}$$
(3.41)

Où w() est une fonction de pondération dans le bloc, p est le pixel manquant à cause de la saturation dans un bloc de l'image courante,  $q_k$  sont les 16 pixels candidats pour le pixel p,  $MV_k$  sont les vecteurs de mouvement associés aux pixels candidats  $q_k$ ,  $r(p,q_k)$  une fonction qui calcule l'inverse de la distance géométrique entre le pixel p et les pixels  $q_k$ ,  $c(p,q_k)$  est une fonction qui donne l'inverse de la pseudo similarité de la couleur.

La génération HDR est réalisée en utilisant la fonction de Debevec et Malik [Debevec et al., 1997] pour la fusion des différentes irradiances des différentes images LDR.

## 3.3.4.2/ MÉTHODE UTILISANT UN ALGORITHME D'OPTIMISATION BASÉ SUR UN PATCH TEMPORELLEMENT COHÉRENT

Cette méthode [Kalantari et al., 2013] utilise une synthèse par patch afin d'améliorer la robustesse dans la détection du flux du mouvement. Selon Kalantari et al [Kalantari et al., 2013], l'algorithme d'optimisation basé sur des patchs temporellement cohérents peut produire une vidéo HDR de haute qualité à partir d'une séquence d'images LDR capturée avec une caméra standard avec différents temps d'exposition. Cette méthode présente l'utilisation du flux optique en combinaison avec une méthode basée sur des patchs similaire à celle présenté par Sen et al [Sen et al., 2012], et ceci pour obtenir une fluidité du mouvement. La combinaison de ces deux techniques permet une robustesse aux défaillances du flux optique dans les zones contenant des mouvements rapides et des occlusions. La méthode basée sur des patch synthétise des textures plausibles et les artefacts sont confinés à de très petites régions proches des limites du mouvement. La combinaison des deux méthode (flux optique et méthode basée sur des patchs) est améliorée par une autre méthode pour calculer des fenêtres de recherche qui varient dans l'espace pour gérer des mouvements complexes. Dans cette méthode, une estimation d'un mouvement initial approximatif est d'abord réalisée, puis utilisée pour calculer une taille de fenêtre de recherche locale. L'algorithme minimise ensuite une fonction d'énergie 3.42 pour la reconstruction de la vidéo HDR en utilisant un algorithme itératif qui itère entre les deux étapes précédentes jusqu'à la convergence. Cette méthode reconstruit les images LDR manquantes qui sont à leurs tour combinées afin de générer les images HDR.

$$E(L_{m,n}) = \sum_{n=1}^{N} \sum_{p \in pixels} [\alpha_{ref,n_{(p)}} \cdot (h(L_{ref,n})_{(p)} - H_{n(p)})^{2} + (1 - \alpha_{ref,n_{(p)}}) \cdot \sum_{m=1,m\neq ref}^{M} \wedge (L_{m,n})(h(L_{m,n})_{(p)} - H_{n(p)})^{2} + (1 - \alpha_{ref,n_{(p)}}) \cdot \sum_{m=1}^{M} TBDS(L_{m,n}, L_{m,n-1}, L_{m,n+1})]$$
(3.42)

Où  $h(L_{ref,n})$  est une fonction qui fait la correspondance de l'image LDR  $L_{ref,n}$  au domaine de radiance linéaire,  $\alpha_{ref,n}$  est une fonction qui donne une approximation de l'exposition d'un pixel de  $L_{ref,n}$ . Ce terme garantit que la reconstruction HDR  $H_n$  soit similaire à  $h(L_{ref,n})$  dans un sens défini par  $L_2$  dans les régions bien exposées. Le second terme de l'équation 3.42 garantit que toutes les images LDR dans une trame sont similaires à l'image HDR dans cette trame et ceci dans le sens défini par  $L_2$  pour les régions qui ne sont pas bien exposées dans l'image de référence. Ce terme maintient la relation entre l'image HDR et les images LDR qui la composent en utilisant une pondération par la fonction chapeau  $\land$ () utilisée pour fusionner les différentes images LDR dans la méthode de Debevec et Malik [Debevec et al., 1997]. Le dernier terme renforce la cohérence temporelle, il est inspiré de la méthode proposée par Shechtman et al [Shechtman et al., 2010]. *TBDS* (acronyme anglais de Temporal BiDirectional Similarité bidirectionnelle de l'image LDR  $L_{m,n}$  à ses équivalents dans les trames précédentes (Lm, n - 1) et suivantes(Lm, n + 1).

$$TBDS(L_{m,n}, L_{m,n-1}, L_{m,n+1}) = BDS(L_{m,n}, L_{m,n-1}) + BDS(L_{m,n}, L_{m,n+1}).$$
(3.43)

Où la similarité bidirectionnelle basé sur un patch BDS est la métrique proposée par Simankov et al [Simakov et al., 2008]. avec les contraintes  $f_S^T(p)$  et  $w_S^T(p)$  dans la recherche basée sur l'estimation du mouvement local pour améliorer la cohérence temporelle donnée en équation 3.44.

$$BDS(T,S) = \frac{1}{|S|} \sum_{p \in pixels} \min_{i \in f_S^T(p) \pm w_S^T(p)} D(s(p), t(i)) + \frac{1}{|T|} \sum_{p \in pixels} \min_{i \in f_S^T(p) \pm w_S^T(p)} D(t(p), s(i))$$
(3.44)

Où s(p) et t(p) représente le patch centré au pixel p dans l'image source et l'image cible, D() est la somme des différences au carré (SSD) entre deux patchs,  $f_S^T(p)$  est l'approximation de flux de mouvement au pixel p de S à T et  $w_S^T(p)$  sert à mettre à l'échelle la fenêtre de la recherche autour de ce pixel.

Cette méthode consiste en trois étapes principales :

- **1. Estimation initiale du mouvement :** Un mouvement brut est estimé dans les deux directions entre les trames consécutives par  $f_T^s(p)$  et  $f_S^T(p)$  dans l'équation 3.44 . Un modèle planaire est utilisé (transformée de similarité) pour le mouvement global et l'estimation du mouvement local est réalisé en utilisant le flux optique.
- **2. Calcul de la fenêtre de recherche :** Une taille de fenêtre est calculée pour chaque vecteur de flux  $(w_T^s(p)$  et  $w_S^T(p)$  dans l'équation 3.43 ). Cette carte de fenêtre de recherche est utilisée comme la taille de la fenêtre de recherche autour de chaque estimation initiale du mouvement.
- 3. Reconstruction vidéo HDR : Une méthode itérative en deux étapes est utilisée pour minimiser l'équation 3.42. La première étape consiste à une recherche et sélection de patch contraint multi-échelles pour minimiser le dernier terme de l'équation 3.42. La deuxième étape (étape de fusion et génération HDR) minimise les deux premiers termes de l'équation 3.42 en utilisant la méthode d'injection de référence [Sen et al., 2012]. L'algorithme itère entre les deux étapes jusqu'à la convergence, ce qui permet de reconstruire les images LDR manquantes et produire les images HDR finales.

Une classification complète des différentes méthodes de suppression et de correction de ghost détaillées ci-dessus, est donnée dans la figure 3.6. C'est une mise à jour de la classification donnée dans l'état de l'art [Srikantha et al., 2012b]. Cette classification est réalisée selon la majorité des paramètres et propriétés des différentes méthodes (domaine de fusion, utilisation de carte de ghost, paramètres intrinsèques des méthode sont déduits automatiquement ou donnés manuellement par l'utilisateur, suppression de tout les objets en mouvement ou fixation d'une occurrence de celui-ci et nombre d'expositions LDR utilisées). Comme on peut le constater sur la figure 3.6, les méthodes peuvent passer d'une classe à l'autre selon le critère de classement. Prenant l'exemple de la méthode d'ordre des pixels et celle de multi-seuils (seuillage à multi-niveaux) lorsque le critère est l'utilisation de la carte de ghost, ces deux méthodes appartiennent à la même classe. Cependant, elles appartiennent à deux classes différentes lorsque elles sont classées selon la fixation d'objets en mouvement. Ainsi, une méthode peut appartenir à différentes classe en même temps même si ces dernières sont complètement opposées (la méthode basée sur le gradient appartient à la classe des méthodes qui suppriment complètement l'objet en mouvement et elle appartient en même temps à la classe de celles qui gardent une occurrence d'objets en mouvement) offrant ainsi une flexibilité d'utilisation selon le besoin.

# 3.4/ Systèmes matériels pour la vidéo HDR

Plusieurs systèmes matériels permettant la capture HDR ont été proposés dans l'état de l'art. Ces systèmes utilisent des approches matérielles différentes afin d'augmenter la gamme de dynamique des images capturée et générer un flux vidéo à grande gamme de dynamique en temps-réel.

## 3.4.1/ Système basé sur l'adaptation de la gamme dynamique

Le premier système a été proposé par Nayar et Branzoi. [Nayar et al., 2003]. Pour augmenter la gamme dynamique capturée, Nayar et Branzoi ont utilisé le principe des pixels à exposition adaptative. L'adaptation d'exposition proposée est réalisée dans le domaine optique (pendant la formation de l'image). Cette adaptation tend à changer l'éclairement incident de chaque pixel en se basant sur la radiance de la scène que le pixel est censé mesurer. Ceci est réalisé en utilisant un atténuateur optique contrôlable devant le système optique de la caméra, comme montré dans la figure 3.7. La luminosité *I* mesurée par un pixel (détecteur) est utilisée pour ajuster le facteur de transmission *T* de la cellule atténuatrice correspondante au pixel détecteur.

Le modèle montré dans la figure 3.7 est simple à réaliser, car l'atténuateur à cristaux
Fu	Fusion Domaine de la radiance			-									_
		Domain	aine de l'image										
		réglage des	Manuel	<u> </u>									
D / 1		seuils	Automatique	-									
suppression de ghost		Carte de	Oui	<u> </u>					_				
		ghost	Non	├									
		Nombre	Petit (<4)	<u> </u>									
		d'exposition	Grand (>5)										
Ré	Résultat Fixation d'objets en mouvemen d'une une seule position			$\vdash$									
f	final Suppresion des objets												
	en mouvement		h										
	Méthodes			I 🖡	¥	¥	¥	¥	¥	Ļ	¥	¥	↓
	Variance				X		X		X		X		X
		Entrop	ie		X		X		х		x		X
		Prédicti	on		X		X		х		x		X
		Ordre des p	ixels	X		X			х	X			X
		Multi-se	uils		X		х		х	Х			х
		Bitmap	)		X		X		х		x	X	
		RANSA	С	X	X		X		х		X		X
	Gradient			X	X		X	X		X		X	
	Fonction de pondération				X	X		X		X			X
	Densités de probabilité conjointes				X	X			X		x		X
		Flux opti	que		X		X		X		X		X
	Compensation du mouvement				X		X	X			X		X
	Pat	ch temporelleme	nt cohérent		X		X	х			X		X
	Estin	ation de mouven	nent par bloc		X		X	X			X		X

FIGURE 3.6 – La classification des différentes méthodes de détection et de suppression de ghost selon les critères de classification donné dans[Srikantha et al., 2012b].



FIGURE 3.7 – Le schéma de système HDR proposé par Nayar et Branzoi. [Nayar et al., 2003].

liquide qui module la lumière incidente, est placé entre l'objet (scène) et l'optique de la caméra. Les tâches critiques de cette méthode sont la mise en place de la carte de correspondance entre l'élément atténuateur et le pixel détecteur ainsi que le contrôle en temps-réel du modulateur de la lumière incidente (atténuateur). Le système d'imagerie développé par Nayar et Branzoi. [Nayar et al., 2003] est illustré dans la figure 3.8. Il est composé d'une caméra vidéo couleur, un objectif devant lequel ils ont fixé un panneau LCD et une carte électronique pour le LCD qui leur permet d'entrer leur fonction de transmission en tant que signal vidéo NTSC. La sortie de la caméra vidéo est connectée à une carte qui numérise et capture des images ( $640 \times 480$ ) à 30 fps. L'algorithme de contrôle est exécuté sur une station du travail Dell dotée d'un processeur Pentium 4 et de 1*Go* de DDRAM.



FIGURE 3.8 – La caméra AGD proposée par Nayar et Branzoi. [Nayar et al., 2003].

#### 3.4.2/ Système HDR basé sur plusieurs capteurs d'image

Un système vidéo HDR récent utilisant 4 capteurs d'image LDR est proposé par Kronander [Kronander et al., 2014], le système proposé est similaire à celui proposé par Tocci et al. [Tocci et al., 2011] qui n'utilise que 3 capteurs d'image LDR. Le but de ce système est de générer un flux d'images HDR basé sur des données provenant d'un système d'imagerie constitué de plusieurs capteurs d'image. Ce dernier capture un ensemble de 4 images LDRs non démosaïqués (brutes CFA) à différents temps d'exposition. L'algorithme proposé pour ce système est conçu spécialement pour des systèmes multi-capteurs où les capteurs acquièrent les images de la scène à travers une lentille frontale commune, mais le trajet de la lumière est divisé sur les différents capteurs par un agencement de séparateur de faisceau (miroirs semi-réfléchissants) comme le montre la figure 3.9 à gauche.

Le système proposé par Kronander nécessite une étape de calibration critique dans l'alignement géométrique des différents capteurs d'image et la calibration radiométrique, car bien qu'il utilise des capteurs d'images de même fabriquant et de même référence, leurs courbes de réponse ne sont pas identiques tout en prenant en compte les variations intro-



FIGURE 3.9 – La caméra ADR proposé par Kronander et al. [Kronander et al., 2014].

duites par le système de division de faisceaux lumineux. L'implémentation matérielle du système proposé par Kronander est montré dans la figure 3.9 à droite, il a utilisé 4 capteurs d'image de bonne qualité fabriqués par Kodak (KAI - 04050 CCD ) d'une résolution spatiale de  $2336 \times 1752$  pixels. La caméra fournit un flux d'image de 4 *MP* à 26 *image/s* en implantant une partie de l'algorithme de génération HDR sur le GPU de la Nvidia GForce 680. La dynamique maximale que cette caméra peut fournir est de 24 f - stops l'équivalent de 144 *dB*.

# 3.5/ ARTÉFACTS GÉNÉRÉS PAR LA SUPPRESSION DE GHOST

Le ghost est considéré comme l'artefact principal limitant la génération HDR à partir de plusieurs expositions. De plus, sa correction réduit la gamme de dynamique de l'image HDR résultante et engendre d'autre artefacts.

1. Les différences visuelles

C'est l'apparition de nouveaux détails (objet ou texture) dans l'image HDR alors que ces détails n'existent dans aucune des images LDR contribuant dans la génération HDR. Cela peut être aussi la disparition dans l'image HDR de caractéristiques présentes dans les images LDR.

2. Les artefacts de fusion

Ces artefacts apparaissent lorsqu'un objet en mouvement est fusionné avec l'arrière-plan. Ce type d'artefact est commun pour les algorithmes qui visent à éliminer les objets en mouvement.

3. Les incohérences des gradients

Elles se produisent lorsque l'image HDR contient de nouveaux gradients qui sont absents dans toutes les expositions ou lorsque les gradients qui existent dans les différentes expositions en entrée au processus de génération HDR, n'apparaissent plus dans l'image résultante.

4. Faible gamme dynamique

C'est la perte en contraste dans l'image HDR au niveau des régions contenant des objets en mouvement dans la pile des images LDR. Cette perte en gamme dynamique est liée au fait que certains algorithmes de suppression de ghost choisissent une seule image LDR (généralement celle prise comme image de référence) pour les régions contenant des objets en mouvement.

#### 3.5.1/ MÉTRIQUES POUR LA SUPPRESSION DE GHOST

Afin de comparer les résultats de la génération HDR (en utilisant la technique de multiples expositions) issus des différentes méthodes de correction de ghost, plusieurs métriques ont été proposées dans l'état de l'art. Une étude approfondie des métriques objectives de qualité HDR [Hanhart et al., 2015] a montré que les métriques les plus performantes et adaptées au domaine de la HDR sont celles de HDR-VDP-2. Cependant, HDR-VDP-2 donne une carte de probabilité, n'est pas quantifiable et nécessite une image HDR de référence. On a donc choisi d'autres métriques objectives récentes dédiées à mesurer la qualité de suppression de ghost en imagerie HDR (proposées par [Tursun et al., 2016]. Ces métriques sont quantifiables et elles n'utilisent pas d'image HDR de référence (voir [Tursun et al., 2016]).

#### 3.5.1.1/ MÉTRIQUE DE FUSION

Les artefacts de fusion se produisent lorsque deux pixels ou plus ayant des valeurs d'irradiance différentes sont combinés pour générer un seul pixel HDR. La présence de cet artefact peut être détectée si le poids de contribution de chaque pixel au résultat final est connu. Cette contribution est donnée par les poids  $\omega$  utilisés lors de la création HDR par la fonction  $I(p) = k \sum_{n=1}^{N} \omega_n(p) E_n(p)$ , I est l'image HDR, p représente l'indice des pixels et k représente une constante de normalisation.  $E_n$  est la  $n^{ieme}$  exposition convertie dans le domaine de l'irradiance en utilisant la fonction de réponse inverse de la caméra et en divisant par le temps d'exposition. L'objectif de la métrique de fusion est de détecter

les pixels qui diffèrent en irradiance et ayant des poids élevés. Cependant, les poids utilisés sont généralement inconnus et ils doivent être estimés (voir plus de détails dans [Tursun et al., 2016]).

$$Q_{Ub}(p) = \sum_{n=1}^{N-1} \sum_{m=n+1}^{N} \left( \frac{\tilde{\omega}_n(p) + \tilde{\omega}_m(p)}{2} g(\tilde{\omega}_n(p), \tilde{\omega}_m(p)) h'(E_n(p), E_m(p)) W_{n,m}(p) \right)$$
(3.45)

Où la fonction  $g(\tilde{\omega}_n(p), \tilde{\omega}_m(p)) = 1 - |\tilde{\omega}_n(p) - \tilde{\omega}_m(p)|$  calcule la similarité entre les poids estimés. La fonction h' (voir l'équation 3.46) calcule la distance euclidienne normalisée h(A, B) entre deux vecteurs d'irradiances et la renvoie uniquement si elle est supérieure à un seuil  $\tau$ . Dans le cas contraire, la valeur 0 est utilisée.

$$h'(E_n(p), E_m(p)) = \begin{cases} 0 & \text{Si} & h(E_n(p), E_m(p)) \le \tau \\ h(E_n(p), E_m(p)) & \text{Autrement} \end{cases}$$
(3.46)

Et  $W_{n,m}(p)$  est une fonction chapeau large [Tursun et al., 2016] qui représente la bonne exposition conjointe d'un pixel p pour les expositions n et m.

#### 3.5.1.2/ MÉTRIQUE DE COHÉRENCE DE GRADIENT

Cette métrique prend comme hypothèse qu'une image HDR ne doit pas contenir un gradient qui n'existe pas dans les images LDR fusionnées. De même, s'il y a un gradient dans toutes les expositions, ce gradient devrait exister dans l'image HDR. Cette métrique consiste à repérer les pixels qui ne respectent pas l'hypothèse précédente. Deux sous-métriques composent la métrique de cohérence de gradient, une pour l'amplitude ( voir l'équation 3.47 ) et une autre pour la direction (voir l'équation 3.48).

$$Q_{UG'_{Mag}}(p) = \min_{n} \frac{\left|\frac{\|\nabla E_{n}\|_{2}}{\|\nabla I\|_{2}} \|\nabla I(p)\|_{2} - \|\nabla E_{n}(p)\|_{2}\right|}{\max\left\{\frac{\|\nabla E_{n}\|_{2}}{\|\nabla I\|_{2}} \|\nabla I(p)\|_{2}, \|\nabla E_{n}(p)\|_{2}\right\}}$$
(3.47)

Où  $\nabla$  est le gradient de l'image. Il est calculé en utilisant l'opérateur de Sobel.  $\overline{||\nabla E_n||_2}$  et  $\overline{||\nabla I||_2}$  sont respectivement les valeurs moyennes de  $||\nabla E_n||_2$  et  $||\nabla I||_2$ . La normalisation avec les valeurs moyennes est effectuée pour rendre les amplitudes de gradient de l'image HDR compatibles avec les amplitudes de gradient des images LDR. Le dénominateur assure la normalisation de la métrique dans l'intervalle [0, 1].

$$Q_{UG'_{Dir}} = \min_{n} \frac{|[(\theta_{I}(p) - \theta_{n}(p) + \pi)mod2\pi] - \pi|}{\pi}$$
(3.48)

Où  $Q_{UG_{Dir}}$  est l'angle minimal entre les directions des vecteurs de gradient.  $\theta_I(p)$  et  $\theta_n(p)$  sont respectivement les directions des vecteurs de gradient dans l'image HDR et l'image LDR *n*.

La carte de l'incohérence de gradient finale est calculée en utilisant la fonction donnée en équation 3.49, où  $Q'_G(p)$  prend les valeurs de  $Q_{UG'_{Dir}}$  et  $Q_{UG'_{Mag}}$ . La fonction renvoie zéro si  $Q'_G(p)$  est inférieur ou égale à un seuil  $\tau$  ou la valeur de  $Q'_G(p)$  dans le cas contraire.

$$Q_G(p) = \begin{cases} 0 & \text{Si} & Q'_G(p) \le \tau \\ Q'_G(p) & \text{Autrement} \end{cases}$$
(3.49)

#### 3.5.1.3/ MÉTRIQUE DE DIFFÉRENCE VISUELLE

Cette métrique [Tursun et al., 2016] est une adaptation de la métrique HDR-VDP-2.2 [Narwaria et al., 2015] afin que les différentes expositions  $L_n$  soient compatibles avec l'image HDR *I*. Pour cela, les images LDR et l'image HDR sont mises à l'échelle pour définir leurs valeurs d'irradiance moyennes à l'unité. Ensuite, la probabilité minimale est calculée pour chaque pixel *p* parmi chaque paire (image *HDR*, image *LDR<sub>n</sub>*). Cette probabilité est choisie comme probabilité de détection de différence visuelle pour ce pixel.

$$Q_{UVd}(p) = \min_{n} V'\left(\frac{I(p)}{\overline{I}}, \frac{L_n(p)}{\overline{L_n}}\right)$$
(3.50)

Où *V* est la carte de la probabilité de détection de différence visuelle générée par le HDR-VDP-2.2 [Narwaria et al., 2015].  $\overline{L_n}$  est la valeur de l'irradiance moyenne de l'image LDR  $L_n$  et  $\overline{I}$  est la valeur d'irradiance moyenne de l'image HDR.

#### 3.5.1.4/ MÉTRIQUE DE GAMME DYNAMIQUE

C'est une métrique qui consiste à calculer la gamme dynamique de l'image HDR dans les zones contenant du mouvement. L'équation 3.51 calcule la différence entre les logarithmes des valeurs des irradiances maximales (à 99%) et les logarithmes des valeurs des irradiances minimales (à 1%) et, afin d'éviter les fausses valeurs liées au bruit et à des artefacts de la génération HDR.

$$Q_{DR} = \log_{10}I(p_{99\%}) - \log_{10}I(p_{1\%})$$
(3.51)

#### 3.5.1.5/ MÉTRIQUE UNIFIÉE DE LA QUALITÉ DE LA SUPPRESSION DE GHOST

Elle est introduite dans le but de donner un score unique représentant la qualité de la génération HDR avec correction de l'artefact de ghost [Tursun et al., 2016]. Ce score, défini dans l'équation 3.52, combine les différentes métriques détaillées antérieurement.

$$Q_{UDQM} = [Q_{Ub} Q_{UG'_{Dir}}^{s} Q_{UVd} Q_{DR} 1] [w_{Ub} w_{UG'_{Dir}} w_{UVd} w_{DR} w_{C}]^{T}$$
(3.52)

Où  $w_i$  indique le poids du score de la métrique *i*, et  $w_c$  est le poids du terme constant qui sera utilisé lors de l'analyse de régression. Les différents poids sont calculés en utilisant la méthode de *recuit simulé adaptatif* (nom anglais : Adaptive simulated annealing "ASA"), plus de détails sur cette dernière sont donnés dans la référence [Ingber, 2000].

Avantages	Inconvénients			
<ol> <li>Métriques conçues spatialement pour quantifier la qualité de sup- pression de ghost en imagerie HDR.</li> </ol>	<ol> <li>Ces métriques prennent en en- trée l'image HDR et les différentes images LDR qui ont contribué à sa</li> </ol>			
<ol> <li>Simples à utiliser car elles n'exigent pas une image HDR de référence.</li> </ol>	génération. L'utilisation de images LDR rend ces métriques obsolètes			
<ol> <li>Quantifiables (elles donnent des score partiels quantifiables et un score unifie qui quantifie la qualité HDR globale).</li> </ol>	dans la quantification de la qualité des images HDR dont le flux LDR est inconnu .			

Tableau 3.1 – Les avantages et inconvénients des métriques objective de suppression de ghost en imagerie HDR

# 3.6/ CONCLUSION

Le problème de ghost dans l'imagerie à haute gamme de dynamique est présenté de manière exhaustif dans ce chapitre. On a commencé par sa définition et les causes de son apparition. Ensuite, les différentes méthodes largement citées dans l'état de l'art de l'imagerie HDR afin de supprimer ou limiter l'artefact de ghost. Chacune de ces méthodes est décrite en détail dans la détection et suppression de ghost. Les méthodes présentées sont classifiées selon leurs utilisations d'images de référence pendant leurs intégration au processus de la génération HDR avec une classe des méthodes qui ont été présentées spécialement pour la vidéo HDR. Cette classification est choisie en prenant en compte la cible matérielle de notre projet qui sera un système de vision basé sur un circuit FPGA avec une contrainte de temps-réel. Cette cible et la contrainte de temps-réel ne tolèrent pas l'utilisation de plusieurs images de référence pour la génération HDR de l'objet en mouvement car ceci est itératif et il exige plusieurs accès à la mémoire. Néanmoins, une autre classification plus large selon les différentes propriétés des algorithmes de correction de ghost est donnée afin d'avoir un aperçu global sur les différentes méthodes de détection et correction de ghost. En plus des méthodes de détection et de suppression de l'artefact de ghost, on a présenté les différents systèmes matériels permettant la génération de la vidéo HDR sans l'artefact de ghost par l'utilisation de plusieurs capteurs d'images tout en gardant le même principe de la fusion de différentes expositions pour la génération d'une seule image HDR. Comme on a présenté les différents artefacts qui peuvent apparaître après l'intégration de l'étape de détection et de suppression du ghost. La qualité des images HDR générées en incluant l'étape de suppression de ghost peut être quantifiée et estimée en utilisant les métriques dédié à la mesure de la qualité de la suppression de ghost présentées en fin de ce chapitre.

Après avoir cerné le problème de ghost dans l'imagerie HDR, le chapitre suivant sera consacré à la pré-sélection des méthodes compatibles avec l'architecture matérielle de système vidéo HDR à mettre en place. Les méthodes pré-sélectionnées seront ensuite implémentées en logiciel sur MATLAB en vue d'une comparaison afin de sélectionner la méthode la plus efficace et adaptée à l'architecture matérielle proposée.

# 4

# ÉVALUATION LOGICIELLE D'ALGORITHMES DE SUPPRESSION DE GHOST EN VUE D'UNE IMPLANTATION MATÉRIELLE

#### Sommaire

4.1	introd	luction	<b>68</b>
4.2	Le log	giciel GhostEva	<b>68</b>
4.3	Test e	et évaluations des méthodes de deghosting	71
	4.3.1	Évaluation de la méthode basée sur l'ordre des pixels	72
	4.3.2	Évaluation de la méthode basée sur la MTB	75
	4.3.3	Évaluation de la méthode basée sur la prédiction	80
	4.3.4	Évaluation de la méthode basée sur la variance	84
	4.3.5	Évaluation de la méthode basée sur la fonction de pondération	
		simple originale	87
	4.3.6	Évaluation de la méthode basée sur la fonction de pondération	
		simple Améliorée	89
4.4	Comp	paraison des résultats des différentes méthodes de suppres-	
	sion o	de ghost	91
	4.4.1	Comparaison des résultats obtenus pour la scène du bureau	92
	4.4.2	Comparaison des résultats obtenu pour la scène du jardin	93
	4.4.3	Comparaison des résultats obtenu pour la scène du cycliste	93
	4.4.4	Comparaison des résultats moyens des trois scènes	95
4.5	Conc	lusion	96

# 4.1/ INTRODUCTION

Dans le chapitre précédent, on a présenté un grand nombre d'algorithmes de détection et suppression d'artefacts de type ghost en imagerie HDR. Néanmoins seulement un nombre limité de ces algorithmes peut prétendre à une implantation matérielle et satisfaire les contraintes temps-réel d'une vidéo HDR. Notre objectif est donc de proposer une implantation matérielle d'une méthode permettant de détecter et de supprimer l'artefact de type ghost en temps réel (60 images/s à une résolution de 1024 × 1280). Les algorithmes retenus pour une étude comparative sont les méthodes qui reposent sur des calculs mathématiques moins complexes et qui ne présentent pas de récursivité. Les algorithmes sélectionnés sont les méthodes basées sur : l'ordre des pixels, une détection et suppression de l'artefact par une seule fonction de pondération, la MTB, la prédiction et la variance. Ces méthodes seront implémentées telles que définies par leurs auteurs dans un premier temps et elle sont ensuite modifiées soit pour améliorer leurs efficacité en détection et de suppression d'artefact de type ghost soit pour les adapter en vue d'une implantation matérielle.

Au préalable, le chapitre débute par la présentation du logiciel GhostEva. Ce logiciel, développé au cours de cette thèse, rassemble les algorithmes de génération HDR, plusieurs méthodes de tone-mapping et intègre une grande variété des méthodes de déghosting décrites précédemment. Leurs versions simplifiées sont aussi présentes afin de sélectionner l'algorithme final destiné à l'implantation matérielle. Dans la continuité de cette description, les résultats d'une comparaison de l'efficacité de correction des artefacts de type ghost est réalisée pour toutes les méthodes sélectionnées afin de proposer le choix de la méthode à implanter matériellement. Comme précisé précédemment, des critères de régularité de l'algorithme, de parallélisme de données ainsi que de complexité des opérateurs nécessaires sont aussi considérées pour réaliser cette sélection.

# 4.2/ LE LOGICIEL GHOSTEVA

GhostEva : acronyme anglais de "Ghost Evaluation lab ", est un outil que nous avons développé en langage Matlab et C. Il permet de tester le flot de génération d'images

HDR en variant les méthodes de génération HDR, les opérateurs de tone-mapping et les méthodes de suppression de ghost. Le logiciel, dont l'interface est présentée figure 4.1, permet de visualiser les résultats, mais aussi de les analyser et les sauvegarder. Les différentes fonctionnalités de notre logiciel GhostEva sont regroupées dans des menus déroulants.



FIGURE 4.1 – L'interface graphique de logiciel GhostEva

 Méthodes de génération HDR : B.1 Il contient la liste des méthodes de génération HDR. Deux méthodes sont implémentées : la méthode de génération HDR dans le domaine de la radiance 2.3.1 proposée par Debevec, et celle de fusion d'expositions dans le domaine de l'image. Pour chacune des méthodes, il est possible de modifier les paramètres tels que le nombre de points utilisés lors de l'extraction de la fonction de réponse de capteur, le paramètre lamda (poids de lissage de la courbe de régression lors de calcule de la fonction inverse de capteur), etc).

- 2. Méthodes de tone-mapping : B.2 Un ensemble de 7 méthodes de tone-mapping a été développé et intégré dans le logiciel GhostEva(Reinhard globale et locale, Reinhard globale, Reinhard globale vidéo, Drago, Duan, Schlick et TumblinRushmeier) et chacune de ces méthodes est paramétrable.
- 3. Méthodes de deghosting : B.3 Un ensemble de 8 méthodes de suppression de ghost est intégré dans le logiciel GhostEva (les versions adaptées ou améliorées des différentes méthodes sont aussi inclues dans ce panel). L'utilisateur peut ainsi choisir la méthode de correction de ghost à utiliser, et l'outil l'insère automatiquement dans le flot de génération HDR. L'utilisateur peut également choisir manuel-lement l'image de référence à utiliser dans l'étape de détection et suppression de ghost (dans le cas contraire l'outil prend automatiquement l'image médiane comme image de référence). Les éventuels paramètres des différentes méthodes de suppression de ghost sont aussi accessible à partir de ce panel (utilisation des opérateurs de morphologie comme la dilatation et érosion, spécifier la taille de noyau de l'opérateur de dilatation et/ou de l'érosion).
- 4. Méthodes d'analyse d'images : B.4 Les différentes mesures de qualité HDR exploitant les métriques de domaine HDR : DQM (Ub, UGDir, UGMag, UVd et DR) et HDR-VDP 2, et les métriques de l'imagerie classiques (SSIM, PSNR, MSE, etc) ont été intégrées. Pour les métriques qui le nécessite (typiquement PSNR), une image de vérité de terrain ou de référence peut aussi être sélectionnée.
- 5. Affichage des résultats : B.5 Cela regroupe les fonctions de visualisation des résultats : image HDR finale, cartes de ghost des différents canaux de couleur (lorsque la méthode sélectionnée lors de la génération HDR utilise une carte de ghost), fonction de réponse du capteur, histogrammes des images HDR (en radiance et tone mappée) ainsi que les résultats des calculs intermédiaires pour certains algorithmes (images de variance, courbe G, etc).
- 6. Sauvegarde des résultats : B.6 L'image HDR tone mappée (un large choix du for-

mat d'enregistrement est disponible), l'image HDR en dynamique complète et résultats d'analyse de l'image au format Excel permet d'être exportés et enregistrés pour une exploitation ultérieure.

# 4.3/ TEST ET ÉVALUATIONS DES MÉTHODES DE DEGHOSTING

L'évaluation logicielle des différentes méthodes de suppression de l'artefact de ghost a été réalisée en utilisant le logiciel GhostEva présenté dans la section précédente. Pour chacune des méthodes pré-sélectionnées, nous réalisons tout d'abord le test de la méthode telle qu'elle a été définie par son auteur (tout en choisissant les paramètres permettant d'obtenir le meilleur résultat), puis dans un second temps nous présentons le test d'une version modifiée par nos soins (soit afin d'améliorer la correction, soit pour diminuer sa complexité algorithmique) puisque l'objectif final demeure l'obtention d'un accélérateur matériel destiné à être embarqué au sein d'une caméra HDR. Afin de rendre reproductible les mesures, nous réalisons les tests sur des bases publiques de données LDR présentés en annexe A. La comparaison des résultats pour chacune des méthodes est réalisée dans un premier temps pour les différentes scènes en considérant : la qualité de la suppression de l'artefact selon la quantité de mouvement (la quantité de mouvement est proportionnelle à la quantité de ghost), la différence de contraste et d'éclairage (pour tester la sensibilité des méthodes de suppression de ghost aux différentes conditions d'illumination), la taille et la nature des objets en mouvement (pour évaluer l'efficacité des méthodes de deghosting à supprimer le ghost quelque soit la taille d'objet en mouvement et la nature de ce mouvement. Car selon l'algorithme utilisé, certaines méthodes peuvent être plus sensibles aux grands objets qu'aux petits et vice-versa ou aux mouvement de translation plus que ceux en rotation et vice-versa). La comparaison a été réalisée par rapport à une génération HDR sans suppression d'artefact de ghost. Le résultat qualitatif sans correction de ghost pour les trois scènes est présenté à titre d'exemple dans la figure 4.2 et une estimation quantitative pour les différentes métriques sélectionnées est proposée dans le tableau 4.1. La génération HDR sans correction de l'artefact de ghost sert de base pour la sélection de la méthode destinée à l'implantation matérielle. Ce choix sera donc réalisé en comparant les différents résultats des méthodes considérées et ceci



pour les trois scènes (voir annexe A).

FIGURE 4.2 – L'image HDR tone mappée générée sans l'étape de suppression de ghost. De haut en bas, dans la première image le ghost est généré par le mouvement de cycliste et de la personne tout à gauche, Dans la deuxième image le ghost est généré par le mouvement de tout les personnes présents dans la scène. Dans la dernière image, le ghost est généré par le déplacement de la boite et du stylo sur le bureau.

## 4.3.1/ ÉVALUATION DE LA MÉTHODE BASÉE SUR L'ORDRE DES PIXELS

L'implémentation logicielle de cette méthode telle qu'elle est présentée par son auteur, a permis d'obtenir les résultats illustrés par la figure 4.3. Comme on peut le voir sur les cartes de ghost obtenues (4.3 à droite où le ghost est représenté par des pixels blanc

Métriques	DR	UVd	UGMag	UGDir	Ub	UDQM
Cycliste	1.14578	0.874142	0.0321053	0.0557083	0.17396	0.320599
Jardin	0.56634	0.976901	0.0250245	0.0471046	0.0423235	0.148959
Bureau	1.15754	0.236271	0.0217171	0.0474149	0.0983513	0.389818

Tableau 4.1 – Les valeurs des différentes métriques HDR obtenues pour les différentes séquences sans l'étape de suppression de ghost

dans la carte), les opérations de morphologie mathématique (étape d'érosion et de dilatation) ont rendu les objets en mouvement homogènes dans la carte de ghost, mais en revanche elle a ajouté plusieurs fausses détections dans l'arrière plan des objets en mouvement comme par exemple les tronc des arbres qui contiennent plusieurs pixels en blanc et le ciel qui apparaît complètement blanc sur la carte de ghost de la base d'image de cycliste, et les livres en arrière plan de la base d'image du bureau (le pourcentage des pixels de ghost sont : 31% pour la scène du bureau, 77 % pour le cycliste et 67 % pour la scène de jardin). Ces fausses détections ont augmenté le nombre de pixels considérés comme pixels des objets en mouvement ce qui par conséquence réduit la dynamique dans ces régions car les pixels des objets en mouvement sont exclus de la contribution dans la génération HDR. Les scores des différentes métriques sont améliorés par rapport à la génération HDR sans suppression de ghost et ceci pour les trois scènes (UDQM passe de 0.32 à 0.36, de 0.14 à 0.16 et de 0.38 à 0.39 pour les trois scènes du cycliste, du jardin et du bureau respectivement) (voir tableaux 4.1 et 4.2). La méthode supprime efficacement le ghost dans la scène de cycliste (un seul et gros objet non homogène en mouvement) et celle du bureau (un objet de taille moyen homogène en mouvement ). Cependant quelques résidus demeurent dans la scène du jardin (plusieurs objets non homogènes en mouvement et des différences de contraste et d'éclairage importantes) par conséquent les scores des métriques obtenus pour cette scène sont faibles et presque la moitié de ceux des deux autres scènes.

Afin de limiter les fausses détections introduites par l'étape de dilatation, nous avons modifié la méthode en désactivant les opérateurs de morphologie mathématique. Afin d'améliorer plus encore la correction, nous avons inclus une étape d'exclusion des pixels des zones sous exposées et saturées par un simple seuillage sur la valeur du pixel considéré ( $5 < P_{valide} < 250$ ). Ce seuillage exclut les pixels des zones saturées de processus

CHAPITRE 4. ÉVALUATION LOGICIELLE D'ALGORITHMES DE SUPPRESSION DE GHOST EN VUE D'UNE IMPLANTATION MATÉRIELLE



FIGURE 4.3 – L'image HDR tone mappée générée en utilisant la méthode de déghosting basée sur l'ordre des pixels avec une étape de dilatation sur la carte de ghost.

Métriques	DR	UVd	UGMag	UGDir	Ub	UDQM
Cycliste	1.15822	0.775647	0.0225213	0.0523203	0.104123	0.366874
Jardin	0.629848	0.940292	0.0391565	0.0605582	0.0425218	0.16359
Bureau	1.16421	0.303445	0.0162107	0.0485831	0.100563	0.393466

Tableau 4.2 – Les valeurs des différentes métriques HDR des résultats obtenus par la méthode utilisant l'ordre des pixels pour la suppression de ghost en utilisant une dilatation avec un noyau de  $3 \times 3$  pixels

de deghosting, mais assure une génération HDR en utilisant tous les pixels des différents images LDR dans ces zones. En plus de la limitation des fausses détections illustré par la diminution des pourcentage des pixels de ghost après modification de la méthode (8% pour la scène du bureau avec une diminution de 23 %, 32 % pour le cycliste avec une diminution de 45 % et 23 % pour la scène de jardin avec une diminution de 44%), les modifications introduites sur l'algorithme basé sur l'ordre des pixels, réduisent le coût de calcul et la mémoire nécessaire tout en simplifiant son implantation matérielle car les opérateurs de morphologie mathématique travaillent par voisinage ce qui nécessite de sauvegarder plusieurs lignes de l'image avant que l'opérateur commence à traiter les pixels du flux entrant (selon la taille de noyau de l'opérateur de morphologie mathématique utilisé). Les résultats obtenus sont donnés dans la figure 4.4.

Comme on peut le visualiser sur la figure 4.4 (à droite), il n'y a que les pixels qui appartiennent aux objets en mouvement qui sont blancs (à 1) et l'arrière plan statique est noir (0). Les modifications mentionnées précédemment ont augmenté la dynamique de l'image finale donnée par le score DR = 1.19 dans le tableau 4.3 pour la scène de cycliste et DR = 1.17 pour la scène de bureau. Par contre la modification introduite a diminué légèrement la dynamique dans la dernière scène de 0.01 . Les scores globaux ont tous augmenté (UDQM a augmenté de 0.36 à 0.40 pour la première scène, de 0.163 à 0.169 pour la scène du milieu et de 0.393 à 0.396 pour la dernière scène). Les résidus de ghost existent toujours dans la deuxième scène (jardin) car cette dernière contient beaucoup d'objets en mouvement et un fort contraste.

Métriques	DR	UVd	UGMag	UGDir	Ub	UDQM
Cycliste	1.19641	0.726168	0.0150705	0.0435779	0.0733579	0.403298
Jardin	0.618082	0.935744	0.0246425	0.0479138	0.0461825	0.169662
Bureau	1.17778	0.294607	0.0137001	0.0446014	0.111706	0.396576

Tableau 4.3 – Les valeurs des différentes métriques HDR des résultats obtenus par la méthode utilisant l'ordre des pixels pour la suppression de ghost sans opération de morphologie mathématique

### 4.3.2/ ÉVALUATION DE LA MÉTHODE BASÉE SUR LA MTB

Les résultats de l'implémentation de cette méthode décrite en section 3.3.2.4, sont donnés dans la figure 4.5. Cette méthode utilise des opérateurs de morphologie mathématique tout comme la méthode basée sur l'ordre des pixels. Afin d'obtenir des résultats



FIGURE 4.4 – L'image HDR tone mappée générée en utilisant la méthode de déghosting basée sur l'ordre des pixels sans dilatation et sa carte de ghost détectée.

optimums, nous avons choisi respectivement des noyaux de tailles  $2 \times 2$  pixels et  $3 \times 3$  pixels pour les masques d'érosion et de dilatation. Le choix des tailles des noyaux est réalisé de manière empirique en testant plusieurs tailles différentes des noyaux avec une comparaison de rendu visuel des résultats.

La carte de ghost sur la figure (4.5 à droite) montre que cette méthode est très limitée car on voit clairement que les objets en mouvement (cycliste et feuilles d'arbres, personnes qui marchent et boite sur le bureau) ne peuvent être distingués de l'arrière plan et du reste des objets de la scène. On observe donc des fausses détections dans les ré-



FIGURE 4.5 – L'image HDR tone mappée générée en utilisant la méthode de déghosting basée sur le MTB et une érosion et dilatation pour sa carte de ghost.

sultats des différentes scènes. Par exemple, dans la scène du cycliste, elles sont visibles au niveau du ciel et de la route, alors dans la scène du bureau, on note des erreurs au niveau du bureau et la bibliothèque. Ces éléments sont normalement fixes et non mobiles. Des erreurs similaires sont observées dans les résultats de la scène de jardin. Les images à gauche de la figure 4.5 présentent plusieurs artefacts dues à la fusion et des résidus de ghost. Néanmoins le score de la métrique DR indique une dynamique assez élevée dans les trois scènes comme indiqué dans le tableau 4.4. En effet, la méthode détecte peu de pixels appartenant réellement aux objets en mouvement mais beaucoup de fausses détection, ce qui explique le nombre élevé des pixels blanc dans les différentes carte de ghost (la première scène présente 51% de ghost, 63% pour la deuxième scène et 61% de ghost dans la dernière scène). Les pixels considérés comme en mouvement apparaissent répartis aléatoirement sur la carte de ghost extraite. Par conséquent, tous les pixels des différentes images LDR contribuent à la fusion et la génération de l'image HDR finale dans les zones statique dont les pixels sont répartis aléatoirement (ce qui donne une dynamique proche de celle obtenue sans suppression de ghost ), mais les valeurs des pixels appartenant aux zones du mouvement sont insérés directement à partir de l'image de référence, ce qui peut introduire des sauts de dynamique suite à l'existence des écarts élevés entre les valeurs de ces pixels et celles des pixels HDR. Le score global de la qualité HDR de la scène du jardin et celle du bureau est faible par rapport à celui de la génération sans la suppression de ghost, ce qui montre une diminution de la qualité HDR lors de l'utilisation de cette méthode.

Métriques	DR	UVd	UGMag	UGDir	Ub	UDQM
Cycliste	1.20313	0.84376	0.0293215	0.0615609	0.0768234	0.388056
Jardin	0.5713	0.960879	0.0387578	0.0669459	0.0447743	0.138761
Bureau	1.16748	0.285187	0.0290205	0.058594	0.100773	0.384672

Tableau 4.4 – Les valeurs des différentes métriques HDR des résultats obtenus par la méthode basée sur le MTB pour la suppression de ghost avec une érosion (noyau  $2 \times 2$  pixels) et une dilatation (noyau  $3 \times 3$  pixels)

La désactivation des opérateurs de morphologie mathématique et l'utilisation de la seule MTB, induit une augmentation du nombre de fausses détections (la scène de cycliste présente 89% de ghost, 85% pour la scène du jardin et 88% de ghost dans la scène du bureau, voir figure 4.6). Dans ce cas, la majorité des pixels appartenant aux cartes de ghost extraites sont considérés comme en mouvement (les cartes de ghost sont majoritairement blanches). Par conséquent, cela se traduit par l'utilisation d'une seule image LDR (l'image de référence) dans la génération HDR. La contribution des autres images LDR apparaît comme du bruit dans l'image HDR finale.

Les résultats de cette méthode contiennent plusieurs artefacts liés aux différences visuelles et aux erreurs de fusion. Les différences visuelles sont illustrées par le score *UVd* dans le tableau 4.5. Les valeurs de ce score ont augmenté pour 2 scènes sur les 3 (pour



FIGURE 4.6 – L'image HDR tone mappée générée en utilisant la méthode de déghosting basée sur le MTB sans opération de morphologie mathématique sur sa carte de ghost.

Métriques	DR	UVd	UGMag	UGDir	Ub	UDQM
Cycliste	1.14624	0.862411	0.0223385	0.0533445	0.080323	0.369187
Jardin	0.602421	0.971667	0.0282862	0.0484934	0.0468448	0.15886
Bureau	1.15177	0.0793849	0.0114313	0.0415624	0.109923	0.396901

Tableau 4.5 – Les valeurs des différentes métriques HDR des résultats obtenus par la méthode basée sur le MTB pour la suppression de ghost sans opération de morphologie mathématique

les métriques UVd, Ub, UGMag et UGDir plus la valeur est petite plus le résultat est bon).

Les erreurs de fusion données par le score de la métrique Ub ont aussi augmenté et ceci

pour les 3 scènes. En plus des artefacts liés aux erreurs de fusion et aux différences visuelles, on constate une chute de la gamme dynamique de 1.20 à 1.14 pour la scène du cycliste et de 1.16 à 1.15 pour la scène du bureau. La qualité HDR finale des résultats obtenus avec la méthode basée sur le MTB sans opération de morphologie mathématique a chuté pour la première et la troisième scène. Bien que les scores globaux *UDQM* des résultats de deux scènes (jardin et bureau) soient améliorés par rapport à l'ancienne version. les images obtenues avec cette méthode ne sont pas réellement HDR. On peut affirmer qu'elle sont équivalentes aux images de référence qui sont prédominantes dans la génération HDR.

# 4.3.3/ ÉVALUATION DE LA MÉTHODE BASÉE SUR LA PRÉDICTION

L'implémentation directe de cette méthode comme elle a été proposée par son auteur présente plusieurs limitations. La carte de ghost extraite par cette méthode est blanche (du mouvement partout dans l'image où les pourcentages de ghost sont 96.5% pour la première scène, 96.7% pour la deuxième scène et 99.8% pour la dernière scène) en raison des fausses détections. Ces dernières sont liées aux estimations des deux fonctions f()et  $f^{-1}()$ , représentant respectivement la réponse de la caméra et sa fonction inverse. On a utilisé une approximation polynomiale à 4 coefficients pour estimer la fonction de capteur. Cette fonction permet d'estimer la valeur d'un pixel à partir d'une valeur d'irradiance. Sachant que l'irradiance est aussi estimée par la fonction inverse de la caméra, les erreurs d'estimation s'accumulent et elles influencent la détection des pixels des objets en mouvement.

Métriques	DR	UVd	UGMag	UGDir	Ub	UDQM
Cycliste	1.1519	0.836027	0.00848886	0.0392914	0.0867949	0.381287
Jardin	0.56454	0.956516	0.0202725	0.0392851	0.0545883	0.147842
Bureau	1.13122	0.199139	0.00489925	0.033774	0.093281	0.396941

Tableau 4.6 – Les	s valeurs des o	différentes i	métriques H	HDR des	résultats	obtenus	par	a
méthode basée su	r la prédiction	pour la sup	pression de	e ghost				

Nous avons essayé d'améliorer la méthode par l'augmentation de nombre de coefficients de l'approximation polynomiale dans un premier temps. Cependant, cette modification augmente considérablement le nombre de ressources logiques demandées par



FIGURE 4.7 – L'image HDR tone mappée générée en utilisant la méthode de déghosting basée sur la prédiction et sa carte de ghost.

une éventuelle implantation matérielle de cette méthode (plus de mémoire pour sauvegarder les coefficients, plus d'opérateurs arithmétiques, augmentation de la latence, etc). De plus, l'augmentation du nombre des coefficients de l'approximation polynomiale n'a pas permis d'améliorer le rendement de la méthode. Nous avons décidé de ne pas utiliser la fonction de capteur. Nous avons modifié la méthode afin qu'elle n'utilise que l'estimation de la fonction inverse de la caméra, au lieu de prédire les valeurs des pixels, la méthode modifiée prédit des valeurs de radiance. Donc la prédiction est réalisée dans le domaine de la radiance, où on calcule la valeur de la radiance du pixel P(x, y) à l'exposition  $t_n$ , puis on prédit sa valeur de la radiance à l'exposition  $t_n$  en utilisant les pixels de l'image de référence (voir équation 4.1). Si la valeur absolue de la différence entre la valeur de la radiance de ce pixel et celle prédite est supérieure à un seuil donné, le pixel est considéré comme appartenant à une zone de mouvement (calculé en utilisant l'équation 4.2).

$$\breve{\mathsf{E}}_{ij}^{l} = \frac{\Delta t_l}{\Delta t_k} f_{-1}(Z_{ij}^k) \tag{4.1}$$

Où  $f_{-1}()$  est la fonction de réponse inverse de la caméra,  $\Delta t_l$  est le temps d'exposition qui correspond à l'image des irradiances prédites  $E_l$  et  $\Delta t_k$  est le temps d'exposition correspondant à l'image  $Z_k$  en entrée du processus de prédiction.

$$G_{ij} = \begin{cases} 1 \text{ Si } \left| \breve{\mathsf{E}}_{ij}^k - E_{ij}^k \right| \ge S \text{ euil} \\ 0 \text{ Autrement} \end{cases}$$
(4.2)

où  $E_{ij}^k = \Delta t_k f_{-1}(Z_{ij}^k)$ .

Les régions de saturation dont les valeurs des pixels sont supérieurs à 252 ou inférieurs à 3 sont exclues du processus de prédiction. Ils sont considérés comme appartenant aux zones statiques, car leurs contribution dans la génération HDR est atténuée par la fonction de pondération native de la génération HDR selon l'algorithme de Debevec et Malik.

La version améliorée de la méthode basée sur la prédiction, détecte très bien les zones contenant les objets en mouvement comme on peut le voir sur la carte de ghost de chacune des scènes dans la figure 4.8 à droite (les pourcentage de ghost sont 30% pour la première scène, 12% pour la deuxième scène et 41% pour la dernière scène). Cette méthode est robuste aux faibles changements de luminance, et ceci en utilisant un seuil plus ou moins élevé (dans notre cas on a utilisé un seuil de 0.5). Le choix d'un seuil d'une manière empirique est la seule limitation de cette méthode car il ne peut pas être choisi d'une manière automatique.

Les valeurs des métriques HDR quantifiant la qualité des résultats de la méthode dans le domaine de la radiance sont regroupées dans le tableau 4.7. Les valeurs des différents



FIGURE 4.8 – L'image HDR tone mappée générée en utilisant la méthode de déghosting basée sur la prédiction dans le domaine de la radiance et sa carte de ghost.

Métriques	DR	UVd	UGMag	UGDir	Ub	UDQM
Cycliste	1.15744	0.738018	0.0181323	0.0445136	0.075077	0.384143
Jardin	0.668087	0.928277	0.0201911	0.0446614	0.0458963	0.193616
Bureau	1.2375	0.22617	0.00952681	0.0406711	0.0962102	0.432931

Tableau 4.7 – Les valeurs des différentes métriques HDR des résultats obtenus par la méthode basée sur la prédiction dans le domaine de la radiance pour la suppression de ghost

scores sont améliorées par rapport aux scores de la méthode originale donnée dans le tableau 4.6 et celle de la génération sans suppression de ghost. Les résultats obtenus

traduisent un bon rendu visuel avec une suppression efficace de ghost.

# 4.3.4/ ÉVALUATION DE LA MÉTHODE BASÉE SUR LA VARIANCE

Les résultats visuels de l'implémentation de la méthode basée sur la variance sont illustrés dans la figure 4.9. L'implémentation de l'algorithme décrit en section 3.3.2.1, intègre pour les tests réalisés une dilatation basée sur un noyau de  $(3 \times 3)$  pixels.



FIGURE 4.9 – L'image HDR tone mappée générée en utilisant la méthode de déghosting basée sur la variance avec une étape de dilatation sur sa carte de ghost.

Les cartes de ghost extraites montrent correctement les objets en mouvement dans les trois scènes. L'étape de la dilatation a permis de rendre les zones de mouvement détectées homogènes. Pour la scène du cycliste, les vélos et les personnes sont parfaitement visible en blanc dans la carte de ghost (où le pourcentage de ghost est de 43% pour la scène du cycliste, 27% pour la scène du jardin et 40% pour la scène du bureau). C'est le cas également pour les personnes qui marchent dans la scène de jardin dans la deuxième scène ainsi que pour les boites sur le bureau dans la scène du bureau. La dilatation a supprimé les pixels noir dans les objets en mouvement dans les cartes de ghost et elle a permis aussi d'assurer une continuité des contours des objets en mouvement. Néanmoins, elle étend légèrement la taille des régions en mouvement. Mais cet accroissement impacte peu la qualité des images en sortie, car la méthode basée sur la variance est précise et le nombre de fausses détections est négligeable. Le tableau 4.8 montre les résultats de cette méthode de manière quantitative, où on observe une augmentation des scores globales UDQM des différentes scènes par rapport à la génération sans suppression de l'artefact de ghost. La seule limitation de cette méthode demeure dans la sélection du seuil (le seuil est fixé empiriquement à 0.03 dans l'implémentation ci-dessus).

Métriques	DR	UVd	UGMag	UGDir	Ub	UDQM
Cycliste	1.17726	0.852171	0.0331777	0.0503616	0.152311	0.342438
Jardin	0.643864	0.891714	0.0270481	0.0515465	0.0452518	0.179867
Bureau	1.18752	0.277935	0.0112455	0.0502658	0.105805	0.405404

Tableau 4.8 – Les valeurs des différentes métriques HDR des résultats obtenus par la méthode basée sur la variance pour la suppression de ghost avec une opération de dilatation (noyau de  $3 \times 3$  pixels)

La désactivation de l'opération de la dilatation a permis d'avoir les résultats visuels illustrés sur la figure 4.10. On constate une amélioration de la carte de ghost pour les trois scènes (où le pourcentage de ghost est de 14% pour la scène du cycliste, 6% pour la scène du jardin et 14% pour la scène du bureau). La carte de ghost contient moins de pixels blanc (pixels des régions de mouvement) que lors de l'utilisation d'une étape de dilatation. La suppression de l'étape de dilatation a amélioré la détection dans les zones statiques (moins de fausses détections). En revanche, les pixels des objets en mouvement ne sont pas tous blancs (détectés) dans la carte de ghost extraite. Par conséquent l'objet en mouvement n'est pas homogène dans l'image HDR finale. Les petites tâches

qui apparaissent sur l'objet en mouvement dans l'image HDR (voir le cycliste dans la première scène et la boite sur le bureau dans la dernière scène) correspondent aux pixels non détectés de l'objet en mouvement dans la carte de ghost.



FIGURE 4.10 – L'image HDR tone mappée générée en utilisant la méthode de déghosting basée sur la variance modifiée et sa carte de ghost.

Métriques	DR	UVd	UGMag	UGDir	Ub	UDQM
Cycliste	1.16917	0.842031	0.0284269	0.0506123	0.116425	0.358769
Jardin	0.639082	0.93469	0.0209087	0.0430499	0.0521255	0.178669
Bureau	1.13636	0.153013	0.0114395	0.0429525	0.100053	0.392227

Tableau 4.9 – Les valeurs des différentes métriques HDR des résultats obtenus par la méthode basée sur la variance modifiée pour la suppression de ghost

Le tableau 4.9 montre une diminution légère de la qualité HDR (pour deux scènes sur trois) des résultats obtenus avec la méthode basée sur la variance modifiée par rapport à sa version originale. Cependant tout les scores des différentes métriques sont améliorés par rapport à ceux de la génération HDR sans suppression de ghost.

# 4.3.5/ ÉVALUATION DE LA MÉTHODE BASÉE SUR LA FONCTION DE PONDÉRA-TION SIMPLE ORIGINALE

L'implémentation de la méthode 3.3.2.9 telle qu'elle est définie par son auteur, a permis d'obtenir les résultats illustrés sur la figure 4.11. Les images HDR obtenues montrent que cette méthode ne supprime pas efficacement l'artefact de ghost (des résidus de l'artefact de ghost sont visibles sur les images HDR). L'impression visuelle est confirmée par les métriques HDR. Dans le tableau 4.10, les scores obtenus sont inférieurs à ceux obtenus sans l'utilisation d'algorithme de suppression de ghost pour la scène de cycliste et celle du bureau. Néanmoins, les scores obtenus pour la scène de jardin sont améliorés par rapport à ceux de la génération sans suppression de ghost. Car pour cette scène, la méthode a détecté une partie des régions du mouvement, cette détection est dû au fort contraste de la scène qui a favorisé d'avantage la détection de la non monotonie de la fonction de réponse de capteur même si le teste a été réalisé sur des valeurs de la fonction inverse de la caméra. L'image de la contribution des différentes images LDR dans la génération HDR illustrée sur la figure 4.11 à droite, montre que pour la scène de jardin dans la zone du mouvement, la génération HDR est réalisée majoritairement par l'image numéro trois seulement. Alors que pour les deux autres scènes, la méthode a utilisé majoritairement l'image de référence (numéro 3) pour la génération de l'image HDR. La contribution des autres images LDR n'est pas nulle, car la méthode utilise une pondération. Cette contribution est visible au niveau des contours. Malgré le faible poids attribué aux pixels des contours, ces derniers restent toujours différenciables de l'arrière plan à moins d'utiliser un poids très faible.

CHAPITRE 4. ÉVALUATION LOGICIELLE D'ALGORITHMES DE SUPPRESSION DE GHOST EN VUE D'UNE IMPLANTATION MATÉRIELLE



FIGURE 4.11 – L'image HDR tone mappée générée en utilisant la méthode de déghosting basée sur la fonction de pondération simple originale.

Métriques	DR	UVd	UGMag	UGDir	Ub	UDQM
Cycliste	1.03851	0.843354	0.019903	0.0516437	0.117359	0.313334
Jardin	0.600494	0.944636	0.0184189	0.0429708	0.04083	0.169826
Bureau	1.13759	0.0693318	0.00955047	0.0393401	0.106765	0.394583

Tableau 4.10 – Les valeurs des différentes métriques HDR des résultats obtenus par la méthode basée sur la fonction de pondération simple originale pour la suppression de ghost

### 4.3.6/ ÉVALUATION DE LA MÉTHODE BASÉE SUR LA FONCTION DE PONDÉRA-TION SIMPLE AMÉLIORÉE

La méthode basée sur la fonction simple de pondération permet d'éviter ou de limiter l'apparition d'éventuelles discontinuités de dynamique qui apparaissent comme des tache sur l'image HDR tone mappée. Il n y a pas de sauts de dynamique qui apparaissent comme des taches sur l'image HDR tone mappée. En effet, cette méthode cherche à adapter le poids des pixels selon leurs valeurs relativement aux pixels correspondant dans l'image de référence. Ainsi, la continuité dans la fonction de pondération est assurée, permettant l'obtention d'une image HDR plus homogène. L'auteur de cette méthode a utilisé les valeurs des pixels en sortie de la fonction inverse de la caméra. Ces valeurs sont dans le domaine de la radiance. Cependant, elles sont systématiquement différentes entre deux images LDR acquises pour différents temps d'exposition. La distance géométrique entre ces deux valeurs dépend plus de l'écart entre les deux temps d'exposition que du mouvement. Donc la distance géométrique entre ces deux valeurs ne peut être utilisée comme une condition discriminatoire du mouvement. Néanmoins, les valeurs à la sortie de la fonction inverse de la caméra (les valeurs X) sont des combinaisons des irradiances E équivalentes des valeurs des pixels de l'image LDR et le temps d'intégration  $\Delta t$  de cette image (  $X = E\Delta t$ ). En se basant sur la monotonie de la fonction inverse de la caméra et en supposant que l'éclairage de la scène est constant, les valeurs d'irradiance E correspondantes sont équivalentes quel que soit le temps d'intégration. Si l'éclairage de la scène est constant, tout changement entre deux irradiances situées sur le même emplacement pixel est donc dû au mouvement. Afin d'améliorer cette méthode, on a remplacé la condition de différenciation en pondérant selon l'écart entre les valeurs de la fonction inverse de la caméra X des différents pixels par l'utilisation de l'écart entre les irradiances E correspondantes de ces pixels 4.3. Le poids obtenu à partir de la fonction de pondération modifiée est utilisé comme un facteur modulateur du poids de la génération HDR natif.

$$w(Z_{ij}^k) = \frac{[a(Z_{ij}^{ref})]^2}{[a(Z_{ij}^{ref})]^2 + [(\frac{f^{-1}(Z_{ij}^{ref})}{\Delta t_k} - \frac{f^{-1}(Z_{ij}^{ref})}{\Delta t_{ref}})/\frac{f^{-1}(Z_{ij}^{ref})}{\Delta t_{ref}}]^2}$$
(4.3)

$$w_{dg}(Z_{ij}^k) = \frac{[a(Z_{ij}^{ref})]^2}{[a(Z_{ij}^{ref})]^2 + [\frac{E_{ij}^k - E_{ij}^{ref}}{E_{ij}^{ref}}]^2}$$
(4.4)

Où  $w(Z_{ij}^k)$  représente la pondération affectée au pixel (i, j) pour une exposition k. f(Z) est la fonction de réponse de la caméra  $\Delta t_k$  le temps d'exposition pour l'image k et  $\Delta t_{ref}$  est le temps d'exposition de l'image de référence. La fonction a(x) a été conservée en préservant la définition donnée en référence [Srikantha et al., 2012b]. La fonction a() est appliquée pour chaque pixel x de l'image LDR de référence, normalisée entre 0 et 1, tel que défini en équation 4.5 :

$$a(x) = \begin{cases} 0.058 + 0.68(x - 0.85) & if x \ge 0.85\\ 0.04 + 0.12(1 - x) & if x < 0.85 \end{cases}$$
(4.5)

La modification introduite sur la fonction de pondération améliore les résultats de détection et ainsi de suppression de ghost. Les images HDR obtenues 4.12 en utilisant cette méthode présentent une haute qualité visuelle (moins de résidus de ghost). Le tableau 4.11 montre que toutes les métriques HDR sont meilleures que celles obtenues lors de la génération HDR sans suppression de ghost et ceci pour les trois scènes. Les images de contribution (voir la figure 4.12 à droite, on a utilisé les images de contribution et non plus les cartes de ghost car cette méthode n'utilise pas de carte de ghost) montrent que toutes les images LDR contribuent à la génération de l'image HDR finale. De plus, seule l'image de référence contribue dans les zones en mouvement.

Métriques	DR	UVd	UGMag	UGDir	Ub	UDQM
Cycliste	1.15802	0.745663	0.0149487	0.0465387	0.0760641	0.386193
Jardin	0.632995	0.938411	0.0203259	0.0416732	0.0473189	0.17868
Bureau	1.16922	0.210566	0.00581236	00.0372794	0.11258	0.402524

Tableau 4.11 – Les valeurs des différentes métriques HDR des résultats obtenus par la méthode basée sur la fonction de pondération simple modifiée pour la suppression de ghost. On présente ici non pas les cartes de ghost à droite (qui n'existent pas pour cette méthode) mais les contributions des images LDRs

Selon le tableau 4.11, la méthode améliore d'autant plus les scores des différences visuelles *UVd* et ceux de la cohérence des gradients *UGMag* et *UGDir* que les scores



FIGURE 4.12 – L'image HDR tone mappée générée en utilisant la méthode de suppression de ghost basée sur la fonction de pondération simple modifiée.

de fusion *Ub* et de la dynamique *DR*. Cette différence est due à la contribution de la fonction de pondération dans la génération HDR de toute l'image et non seulement dans les régions du mouvement.

# 4.4/ COMPARAISON DES RÉSULTATS DES DIFFÉRENTES MÉ-THODES DE SUPPRESSION DE GHOST

Afin de comparer les méthodes testées et évaluées dans la section précédente, on a réorganisé les résultats des métriques HDR des différentes méthodes dans des tableaux

par scène. La comparaison débute par l'analyse des résultats des différentes méthodes pour chacune des scènes, puis, la moyenne des résultats est présentée afin de choisir la méthode finale à implanter en matériel et à intégrer au sein de notre caméra.

#### 4.4.1/ COMPARAISON DES RÉSULTATS OBTENUS POUR LA SCÈNE DU BUREAU

Les résultats des différentes méthodes pour la scène du bureau sont donnés dans le tableau 4.12 et affiché sur la figure 4.13. La méthode basée sur la prédiction dans le domaine de la radiance (prédiction modifiée) donne le meilleur résultat selon la métrique UDQM de la qualité HDR (elle a le score de qualité HDR globale le plus élevé UDQM = 0.43293). Elle est suivie par la méthode basée sur la variance puis la méthode basée sur la fonction de pondération modifiée. Cependant la méthode basée sur l'ordre des pixels modifiée est placée en troisième position selon son rendement en gamme dynamique DR (après la méthode basée sur la prédiction modifiée a un score d'artefact de différences visuelles élevé UVd = 0.29460. Cette limitation a diminué donc son score qualité HDR globale UDQM.

Métriques	DR	UVd	UGMag	UGDir	Ub	UDQM
Sans deghosting	1.15754	0.23627	0.02171	0.04741	0.09835	0.38981
Ordre des pixel	1.16421	0.30344	0.01621	0.04858	0.10056	0.39346
Ordre des pixel modifié	1.17778	0.29460	0.01370	0.04460	0.11170	0.39657
МТВ	1.16748	0.28518	0.02902	0.05859	0.10077	0.38467
MTB modifié	1.15177	0.07938	0.01143	0.04156	0.10992	0.39690
Prédiction	1.13122	0.19913	0.00489	0.03377	0.09328	0.39694
Prédiction modifiée	1.2375	0.22617	0.00952	0.04067	0.09621	0.43293
Variance	1.18752	0.27793	0.01124	0.05026	0.10580	0.40540
Va riance modifiée	1.13636	0.15301	0.01143	0.04295	0.10005	0.39222
WFS originale	1.13759	0.06933	0.00955	0.03934	0.10676	0.39458
WFS modifiée	1.16922	0.21056	0.00581	0.03727	0.11258	0.40252

Tableau 4.12 – Tableau de comparaison des résultats des différentes méthodes pour la scène du bureau



FIGURE 4.13 – Graphique de la comparaison des résultats des différentes méthodes pour la scène du bureau.

#### 4.4.2/ COMPARAISON DES RÉSULTATS OBTENU POUR LA SCÈNE DU JARDIN

La scène de jardin est la scène la plus contraignante. Toutes les méthodes donnent des scores de qualité HDR globale UDQM faibles pour cette scène. Ces résultats sont liés aux contraintes extrêmes de la scène (plusieurs objets en mouvement, tailles différentes des objets en mouvement, dynamique de la scène élevée, différents types de mouvement, texture élevée). Selon les résultats des métriques HDR des différentes méthodes (tableau 4.13), La méthode basée sur la prédiction dans le domaine de la radiance obtient le meilleur score de qualité HDR globale (UDQM = 0.19361). Elle est suivie par la méthode basée sur la variance et celle basée sur la fonction de pondération modifiée (respectivement UDQM = 0.17986 et UDQM = 0.17868).

#### 4.4.3/ COMPARAISON DES RÉSULTATS OBTENU POUR LA SCÈNE DU CYCLISTE

Cette scène est caractérisée par peu de mouvement constitué du mouvement d'un seul objet (le cycliste) et les feuilles d'arbres. De plus, la scène a une gamme dynamique faible. Toutes les méthodes de suppression de ghost enregistrent des scores de qualité globale UDQM élevés. La méthode basée sur l'ordre des pixels modifiée présente le meilleur résultat en terme de qualité HDR globale (dont le score UDQM = 0.40329) suivi

Métriques	DR	UVd	UGMag	UGDir	Ub	UDQM
Sans deghosting	0.56634	0.97690	0.02502	0.04710	0.04232	0.14895
Ordre des pixel	0.62984	0.94029	0.03915	0.06055	0.04252	0.16359
Ordre des pixel odifié	0.61808	0.93574	0.02464	0.04791	0.04618	0.16966
MTB	0.57130	0.96087	0.03875	0.06694	0.04477	0.13876
MTB modifié	0.60242	0.97166	0.02828	0.04849	0.04684	0.15886
Prédiction	0.56454	0.95651	0.02027	0.03928	0.05458	0.14784
Prédiction modifiée	0.66808	0.92827	0.02019	0.04466	0.04589	0.19361
Variance	0.64386	0.89171	0.02704	0.05154	0.04525	0.17986
Variance modifiée	0.63908	0.93469	0.02090	0.04304	0.05212	0.17866
WFS originale	0.60049	0.94463	0.01841	0.04297	0.04083	0.16982
WF S modifiée	0.63299	0.93841	0.02032	0.04167	0.04731	0.17868

Tableau 4.13 – Tableau de comparaison des résultats des différentes méthode pour la scène du jardin



FIGURE 4.14 – Graphique de la comparaison des résultats des différentes méthodes pour la scène du jardin.

par la méthode basée sur la pondération simple modifiée avec un score UDQM = 0.38619(voir le tableau 4.14). On a exclu la méthode basée sur le MTB même si cette dernière a obtenue un score UDQM élevé. L'exclusion de la méthode basée sur le MTB n'est pas liée à une erreur dans le calcul des différentes métriques, mais elle est due à son utilisation d'une seule image LDR (celle de référence) dans la génération HDR. La scène de cycliste est de faible gamme de dynamique et les objets en mouvement sont tous présents dans l'image de référence. L'image HDR idéale de cette scène est très proche de l'image de référence. C'est pour cette raison que la méthode basée sur le MTB (qui

94
n'a utilisé que l'image de référence) a obtenu un bon résultat en métriques de qualité HDR.

Métriques	DR	UVd	UGMag	UGDir	Ub	UDQM
Sans deghosting	1.14578	0.87414	0.03210	0.05570	0.17396	0.32059
Ordre des pixel	1.15822	0.77564	0.02252	0.05232	0.10412	0.36687
Ordre des pixel odifié	1.19641	0.72616	0.01507	0.04357	0.07335	0.40329
MTB	1.20313	0.84376	0.02932	0.06156	0.07682	0.38805
MTB modifié	1.14624	0.86241	0.02233	0.05334	0.08032	0.36918
Prédiction	1.1519	0.836027	0.00848	0.03929	0.08679	0.38128
Prédiction modifiée	1.15744	0.73801	0.01813	0.04451	0.07507	0.38414
Variance	1.17726	0.85217	0.03317	0.05036	0.15231	0.34243
Variance modifiée	1.16917	0.84203	0.02842	0.05061	0.11642	0.35876
WFS originale	1.03851	0.84335	0.01990	0.05164	0.11735	0.31333
WF S modifiée	1.15802	0.74566	0.01494	0.04653	0.07606	0.38619

Tableau 4.14 – Tableau de comparaison des résultats des différentes méthode pour la scène de cycliste



FIGURE 4.15 – Graphique de la comparaison des résultats des différentes méthodes pour la scène du bureau.

#### 4.4.4/ COMPARAISON DES RÉSULTATS MOYENS DES TROIS SCÈNES

Comme on peut voir sur le graphe affiché dans la figure 4.16, la méthode basée sur la prédiction modifiée détient le score *UDQM* le plus élevé suivi par la méthode basée sur l'ordre des pixels modifiée et la méthode de pondération simplifiée. Mais en terme

de score de la gamme dynamique *DR*, la méthode de pondération améliorée détient la deuxième place après la prédiction modifiée. La méthode basée sur la pondération modifié enregistre toujours des scores de qualité de gradient élevé par rapport à la méthode de pondération modifiée.

Métriques	DR	UVd	UGMag	UGDir	Ub	UDQM
Sans deghosting	0.95655	0.69577	0.02628	0.05007	0.10487	0.28645
Ordre des pixel	0.98409	0.6731	0.02596	0.05382	0.08240	0.30797
Ordre des pixel odifié	0.99742	0.6521	0.01780	0.04536	0.07708	0.32317
MTB	0.98063	0.6966	0.03236	0.06236	0.07412	0.30382
MTB modifié	0.96681	0.6378	0.02068	0.04780	0.07903	0.30831
Prédiction	0.94922	0.6638	0.01122	0.03745	0.07822	0.30869
Prédiction modifiée	1.02100	0.6308	0.01595	0.04328	0.07239	0.33689
Variance	1.00288	0.6739	0.02382	0.05072	0.10112	0.30923
Variance modifiée	0.98153	0.6432	0.02025	0.04553	0.08953	0.30988
WFS originale	0.92553	0.6191	0.01595	0.04465	0.08831	0.29258
WF S modifiée	0.98674	0.6315	0.01369	0.04183	0.07865	0.32246

Tableau 4.15 – Tableau de comparaison des résultats moyens des différentes scènes pour toutes les méthodes



FIGURE 4.16 – Graphique de la comparaison des résultats moyens des différentes scène pour toutes les méthode.

#### 4.5/ CONCLUSION

L'implémentation logicielle des différentes méthodes de suppression de l'artefact de ghost, a permis de les étudier et de les tester dans différentes conditions (dynamique de

la scène, quantité de mouvement (distance de déplacement, nombre d'objets en mouvement), nature de mouvement (translation, rotation ou translation et rotation) et nature des objets en mouvement). Comme on a pu le voir, l'efficacité de chaque méthode est différente selon ces conditions. Néanmoins, certaines méthodes conservent une bonne qualité HDR quelle que soit la scène filmée (la méthode basée sur la prédiction dans le domaine de la radiance et la méthode basée sur la fonction de pondération améliorée).

L'implémentation logicielle a permis aussi d'étudier la complexité des algorithmes de suppression de ghost pré-sélectionnés. On a montré que la méthode basée sur l'ordre des pixels apparaît comme la plus simple à implémenter en logiciel et en matériel. Selon les résultats obtenus, les méthodes basées sur la prédiction dans le domaine de la radiance et celle de pondération simple améliorée sont les meilleures candidates à une implantation matérielle et une intégration sur la caméra HDR. La méthode de prédiction dans le domaine de la radiance est performante et simple à intégrer par rapport à celle de pondération améliorée. Néanmoins elle utilise un seuil pour détecter les pixels qui appartiennent aux régions contenant du mouvement, et ce dernier dépend de la nature de la scène et sa dynamique ce qui rend le choix d'un seuil optimal difficile et non automatique. En plus de la difficulté à trouver le seuil optimal, la méthode basée sur la prédiction dans le domaine de la radiance remplace les pixels des régions de mouvement par les pixels de l'image de référence directement, ce qui crée des sauts de niveau de dynamique qui apparaissent comme des non homogénéités dans l'image HDR résultante surtout celle tone mappée. Pour ces raisons, on a choisi d'implanter matériellement la méthode basée sur l'ordre des pixels pour sa simplicité et la méthode basée sur la fonction de pondération pour ses bons résultats en métriques de qualité HDR et aussi son bon rendement visuel. Le chapitre suivant sera consacré à l'implantation matérielle de la caméra HDR intégrant l'une des deux méthodes sélectionnées.

# 5

## IMPLANTATION MATÉRIELLE DE LA SUPPRESSION DE GHOST ET SON INTÉGRATION À LA CAMÉRA HDR

#### Sommaire

5.1	introd	uction .		
5.2	La car	méra HDR		
	5.2.1	Le gestio	onnaire de la mémoire vive MMU	
	5.2.2	Le contr	ôleur de l'interface Ethernet 106	
	5.2.3	Le modu	le de tone mapping	
	5.2.4	La géné	ration HDR	
5.3	Le mo	dule de l	a suppression de ghost 111	
	5.3.1	Suppres	sion de ghost par la méthode basée sur l'ordre des pixels 111	
		5.3.1.1	Résultats de l'implantation globale utilisant l'ordre des	
			pixels 113	
	5.3.2	Suppres	sion de ghost par la fonction de pondération 115	
		5.3.2.1	Implantation matérielle de la fonction de pondération	
		5.3.2.2	L'efficacité de la génération HDR	
		5.3.2.3	L'efficacité en suppression de ghost 118	
		5.3.2.4	Implantation matérielle de la fonction de pondération améliorée utilisant le DL de l'exponentiel	
		5.3.2.5	Le module de la génération HDR utilisant le module de la suppression de ghost	
		5326	Résultate d'acquisition	
	533	Suppres	sion de ghost par la fonction de pondération tabulée	
	0.0.0	5331	Le module de la génération HDR utilisant le module de	
		5.5.5.1	suppression de ghost tabulé	
	5.3.4	Résultat	s obtenus par notre caméra HDR utilisant l'architecture	
		tabulée	de module de suppression de ghost	
		5.3.4.1	Efficacité en génération HDR 132	
		5.3.4.2	Efficacité en préservation de texture 134	
		5.3.4.3	Efficacité en génération HDR et suppression de ghost . 134	
5.4	Concl	usion		

#### 5.1/ INTRODUCTION

De nos jours, il existe peu de systèmes d'acquisition d'images HDR ou de cameras HDR, et ceux existants souffrent de plusieurs limitations : artefacts de mouvement (pour les systèmes utilisant des images LDR espacées dans le temps), génération HDR non temps-réel, cadence image faible, coût élevé (caméras HDR utilisant des capteurs d'images dédiées) et systèmes matériels encombrants (par exemple les systèmes basés sur l'utilisation de plusieurs caméras). Dans ce chapitre, on présentera un système matériel d'acquisition d'images HDR temps-réel basé sur un composant de type FPGA. Ce système utilise la technique de la génération HDR à partir de plusieurs images LDR à différents temps d'expositions. Comme cette technique souffre des artefacts liés aux mouvements d'objets dans la scène, des étapes de détection et de suppression de ces derniers ont été intégrées dans la chaîne de génération temps-réel du flux HDR. Les implantations de l'ensemble des modules matériels de cette chaîne de génération HDR sont décrits dans ce chapitre ainsi que les résultats d'acquisitions obtenus pour les différentes versions présentées.

#### 5.2/ LA CAMÉRA HDR

La caméra HDR proposée est une version améliorée de la caméra HDR-ARtiSt [Lapray et al., 2014]. L'architecture interne du système original a été significativement modifiée ainsi que la plate-forme globale. L'étape d'acquisition a été intégralement conservée : le capteur d'images et son contrôle restent inchangés par rapport à la caméra d'origine. Cependant les modules de traitement, de communication et de contrôle global du système ont été profondément modifiés ou améliorés. La plateforme d'origine intégrant ces différentes fonctionnalités, une carte d'évaluation de type Xilinx ML507, a ainsi été remplacée par une de type ML605 (voir la figure 5.1). Cette plate-forme est basée sur un FPGA Xilinx Virtex 6 (*xc6v1x240t*). Le changement de plateforme a permis l'utilisation d'un FPGA plus récent, ce qui a directement impacté les performances du système. La ML507 est basée sur un FPGA de type Virtex 5 utilisant une technologie CMOS 65nm alors que la ML605 repose sur un composant de type Virtex6 utilisant une technologie 40nm. Le potentiel de traitement est directement accru par l'augmentation des ressources logiques disponibles, mais aussi par l'accroissement de la fréquence maximale de fonctionnement (passage de 500 à 600MHz). Un autre avantage de ce changement de technologie est la diminution de la consommation électrique due circuit de 20% par rapport à l'ancienne technologie. En outre, la plate-forme ML605 intègre 512 Mo de mémoire vive de type DDR3 SDRAM. Elle est utilisée pour sauvegarder la pile d'images LDR requises en accès simultané pendant la génération HDR. L'utilisation de ce type de mémoire vive améliore la vitesse d'accès aux données par rapport à la mémoire vive utilisée par l'ancienne plate-forme (la carte d'évaluation ML507 intègre une mémoire vive de type DDR2). L'augmentation de la bande passante permet d'accroître potentiellement le nombre d'images LDR utilisées pour la génération HDR. Ainsi guatre images LDR peuvent être utilisées au lieu de trois précédemment. La qualité HDR de l'image obtenue peut ainsi être étendue surtout lorsque la scène imagée possède une très grande gamme dynamique. Le démonstrateur décrit par la suite utilisera uniquement trois images LDR. La caméra proposée intègre trois interfaces de communication : une interface SPI en interne utilisée pour contrôler le capteur d'image, et deux autres pour la communication externe, une interface Ethernet pour lier le système d'acquisition (la caméra HDR) à l'application logicielle de contrôle qui s'exécute sur un ordinateur hôte et une interface DVI permettant l'affichage des images LDR ou HDR tone mappées sur un écran conventionnel. Une carte adaptation, présentée figure C.1 C) permet l'insertion du capteur d'images à la carte d'évaluation ML605. Ce capteur couleur (EV76C560), produit par E2V, possède intrinsèquement une dynamique classique limitée à 62dB. Sa résolution spatiale est de 1280x1024 pixels. Ce capteur permet la pré-programmation de quatre temps d'exposition différents, ce qui est un avantage pour la génération HDR utilisant la technique de la fusion de multiples expositions. L'architecture interne du système est entièrement repensée grâce à la flexibilité des composants FPGA et a permis l'insertion du module de correction d'artefacts liés

au mouvement au sein de la chaîne de génération HDR. La correction a lieu en temps réel au rythme de la cadence vidéo maximale du capteur, à savoir 60 images/s. Tous les contrôleurs ont été définis spécifiquement pour les interfaces de communication intégrées à la nouvelle plateforme (mémoire DDR3 SDRAM et interface ethernet). La configuration des phases d'acquisition s'effectue via l'interface Ethernet. L'ensemble des registres de contrôle du capteur sont ainsi configurés via cette interface. Une API spécifique a été développée pour réaliser le contrôle global de la caméra. Elle est déportée et installée sur un ordinateur hôte. L'unité de la gestion de la mémoire vive (MMU : acronyme anglais de Memory Management Unit) a été spécialement conçue pour permettre le stockage de l'image courante en cours d'acquisition et la relecture des autres images LDR nécessaires à la génération HDR. Le fonctionnement de la MMU est détaillé dans sous section section 5.2.1. La transmission des données peut s'effectuer suivant deux modes différents, sélectionnables depuis l'API. Le premier mode permet de transmettre un flux d'images HDR à la cadence de 15 images pleine résolution par seconde. Dans ce cas, chaque pixel est codé en format flottant 32-bits simple précision IEEE754. Le nombre d'images par seconde pourrait être augmentés jusqu'à la fréquence maximale du capteur (60 images/s en pleine résolution) avec l'utilisation d'une interface possédant une bande passante supérieure à 1Gbit/s (typiquement Ethernet 10Gbit/s, PCI express, etc). Le deuxième mode permet l'envoi d'un flux d'image tone mappé codé sur 8 bits fonctionnant à la vitesse de capteur d'image (60images/s) en pleine résolution. Ce mode permet ainsi un affichage des résultats. L'ensemble des modules nécessaires au fonctionnement du système, à savoir pour l'acquisition des images, la génération HDR sans artefacts, les différents contrôleurs et enfin les interfaces de transfert et d'affichage de données, ont été décrits dans le langage VHDL afin d'étendre la portabilité sur d'autres plateformes basées sur des composants de type FPGA.

L'architecture globale est illustrée figure 5.2.



FIGURE 5.1 – La caméra HDR.

#### 5.2.1/ LE GESTIONNAIRE DE LA MÉMOIRE VIVE MMU

Le module en charge de la génération HDR requiert plusieurs images de faible gamme de dynamique. Nous en utiliserons ici trois, acquises avec trois temps d'exposition différents : faible, moyen et haut (voir la figure 5.3). La génération HDR nécessite alors trois pixels situés exactement à la même position dans les trois images à chaque coup d'horloge afin de générer le pixel HDR de la position concernée. Ainsi pour la génération de chaque image HDR, l'image en cours et les deux images acquises antérieurement doivent être accessibles simultanément par le module de la génération HDR. La MMU contrôle les trois flux de donnée pixels [Lapray et al., 2014] et permet un accès circulaire des images dans la mémoire centrale globale (SDRAM DDR3).



FIGURE 5.2 – Hardware architecture of the described HDR camera.

La MMU est chargée des échanges de données entre la mémoire globale externe (SDRAM) et le FPGA. En fonctionnement nominal, à savoir dès que deux images ont été acquises, l'accès en écriture est réalisé pour sauvegarder l'image courante pendant qu'un accès en lecture est réalisé afin de charger les deux images LDR les plus récentes (sauvegardées précédemment) à partir de la mémoire globale. Le transfert de données est réalisé ligne par ligne. Une temporisation de données est assurée en utilisant des blocs mémoire doubles port configurés en mode FIFOs (first in first out : première donnée stockée, première donnée lue). Ces éléments, pouvant fonctionner de manière asynchrone en lecture et écriture, permettent de sauvegarder temporairement les données et assurer un changement de domaine d'horloge (entre l'horloge du capteur d'image et celle de

104



FIGURE 5.3 – Le fonctionnement de l'unité de la gestion de la mémoire globale avec la permutation de trois images LDR.

la mémoire DDR3 SDRAM). La MMU lit les données à partir de la mémoire globale en réalisant des temporisations à nouveau à partir de mémoire FIFO, comme illustré dans la figure 5.4. Les deux lignes (L1) issues de la même localisation, mais pour expositions précédentes Img0 et Img1 sont lues dans la mémoire globale, puis sauvegardées temporairement dans les mémoires caches FIFO1 et FIFO2. Ce transfert est réalisé quand le capteur d'image ne fournit aucune donnée c'est-à-dire quand le signal de donnée valide est désactivé ou bien la phase d'intégration de flux lumineux. Dès que le capteur d'image débute le transfert des données relatives à l'exposition en cours, les trois flux de pixels sont donc disponibles pour le module de suppression d'artefact de ghost et celui de génération HDR. Simultanément, la donnée provenant du capteur d'image est temporisée dans une mémoire d'écriture de type FIFO (FIFO3) afin de la sauvegarder dans la mémoire globale. L'accès aux données est réalisé à la fréquence de la DDR3 SDRAM (fixé à 125MHz avec un bus d'interface de 128 bits) dans le mode burst ce qui permet d'atteindre une bande passante 10 fois plus élevée que la bande passante en lecture de capteur d'image (114MHz avec un bus d'interface de 10 bits). Pour cette architecture, le nombre d'images LDR utilisées pour une génération HDR en temps-réel est donc directement lié



aux caractéristiques de la mémoire globale et la fréquence de capteur d'images utilisé.

FIGURE 5.4 – Détails sur la gestion de flot de donnée réalisé par l'unité de gestion de la mémoire MMU (les mémoires FIFOs sont représentées par des cadres bleus et la DDR3 SDRAM par des cadres rouges ).

L'unité de gestion de mémoire est l'élément clé permettant la préparation des trois flux d'images LDR nécessaire à la génération d'un flux images à grande gamme de dynamique, et ceci à la vitesse maximale permise par le capteur d'images. Ce fonctionnement peut être réalisé pour la résolution maximale de  $1280 \times 1024$  avec une latence de deux images.

#### 5.2.2/ LE CONTRÔLEUR DE L'INTERFACE ETHERNET

La caméra HDR développée utilise l'interface Ethernet 1*Gbit* pour communiquer avec un ordinateur hôte distant. Nous avons développé un contrôleur d'interface Ethernet intégrant le protocole UDP (User Datagram Protocole ) de transport de donnée dont l'architecture est illustrée figure 5.5. Ce protocole est utilisé pour envoyer et recevoir les données de l'ordinateur hôte. Nous avons utilisé le protocole UDP (voir le module UDP dans [Alachiotis et al., 2010]) afin d'offrir une bande passante élevée en transmission de données (plus de 80 *Moctet/s* de données vidéo). Cette bande passante ne peut pas être atteinte en utilisant le protocole de transport de donnée TCP (Transmission Control Protocol ). En effet, le protocole UDP utilise moins de données de contrôle de transmission ainsi son implantation matérielle nécessite moins de ressources logiques. Afin d'assurer une transmission de données fiables via l'interface Ethernet utilisant le protocole UDP, nous avons ajouté quelques données de contrôle dans le champ de données de paquet UDP. Ces dernières sont insérées lors de l'encapsulation d'une ligne ou d'une partie de la ligne via le sous-module "Encapsulation des lignes" présenté figure 5.5 et ceci avant la génération des paquets UDP. Deux modes de transmissions ont été intégrés :



FIGURE 5.5 – Architecture du module de transmission des trames de données via l'interface Ethernet.

— Transmission de flux vidéo tone mappé : dans ce mode, la transmission est réalisée ligne par ligne. Deux champs de contrôle sont utilisés (un champ pour l'identifiant d'image "ID Image " et un autre pour l'identifiant de la ligne "ID Ligne" ) pour chaque ligne comme illustré figure5.6. "ID Ligne" contient l'indexe de la ligne envoyée codé sur deux octets. Ce numéro est nécessaire pour garantir une reconstruction correcte de l'image par l'ordinateur hôte. En effet, l'ordre de réception des paquets n'est pas garanti, mais de plus il s'agit de détecter d'éventuelles pertes de lignes. (ID Ligne correspond au nombre de lignes d'une image pleine résolution du capteur). "ID Image " est le numéro de l'image en cours de transmission, il est codé sur un octet et est utilisé pour réactualiser l'image reçue par l'application et également pour détecter une éventuelle perte d'image.

— Transmission de flux vidéo HDR brut : ce mode de transmission est similaire au premier mode. En plus des deux champs "ID Ligne" et "ID Image", il contient un autre champs nommé "ID paquet". En effet, dans cette configuration, la ligne à transmettre est divisée en 4 paquets équivalents à une taille de 1280 octets.

<u>—</u> Données d'une line (1280 Octets des nivels HDR )	ID	ID	ID
4	Paquet	Image	Ligne

	ID	ID
Données d'une line (1280 Octets des pixels HDR tone mappe ou LDR )	Image	Ligne

FIGURE 5.6 – Trame de transmission des données d'une ligne.

Les données provenant de l'ordinateur hôte sont reçues via le même module de communication Ethernet et décapsulées par le même module de contrôle de protocole UDP. Après ouverture d'un paquet UDP, le sous-module "Extraction de données" décode les données de contrôle envoyées via l'application (sur l'ordinateur hôte) et ensuite sépare les flux de contrôle de flux de données (exemple CMD=changement de temps d'exposition moyen et les données = temps ).

Le sous module "CTRL" est responsable du contrôle de l'encapsulation des lignes et celui de l'extraction des données. Il est également en charge de la génération des deux domaines d'horloge nécessaires à la bonne gestion des FIFO asynchrones. En effet, les lectures et les écritures s'effectuent à des rythmes différents dans le cas de la transmission d'un flux vidéo HDR brut.

Le module de communication via l'interface Ethernet utilise moins de 1% des ressources globales du FPGA. Ce module peut fonctionner à une fréquence maximale de 200 MHz, qui est cependant restreinte à la fréquence de l'horloge de l'interface Ethernet de 125MHz et celle de capteur d'image (114MHz).

Type de composant	Nombre de composants	En %
Registres	1637	1%
LUT	1311	1%
Slice	564	1%
GTXE1	1	5%
IBUFDS_GTXE1	1	8%
TEMAC_SINGLE	1	25%
RAMB36E1	4	1%
RAMB18E1	1	1%

Tableau 5.1 – Ressources utilisées par l'implantation matérielle de module de communication éthernet

#### 5.2.3/ LE MODULE DE TONE MAPPING

Ce module permet d'adapter la dynamique et le format de sortie de la vidéo HDR afin de pouvoir visualiser l'information imagée. En effet une image représentée au format flottant 32 bits n'est pas observable sur une interface d'affichage conventionnelle à gamme dynamique traditionnelle (8 bits par canal de couleur). Les images HDR sont ainsi converties en "compressant" la gamme dynamique afin de pouvoir être affichées. Comme mentionné dans le chapitre 2, il existe deux grandes familles d'algorithme de tone mapping. Les méthodes appartenant à la première famille, que l'on peut qualifier d'approches locales, agissent en considérant uniquement le voisinage de chaque pixel alors que les autres traitent l'image d'une manière globale pour tous les pixels de l'image, comme présenté dans le chapitre 2. Les algorithmes de tone mapping locaux nécessitent des calculs complexes donc plus de ressources logiques et de temps d'exécution. Leur implantation apparaît incompatible avec les contraintes de traitement temps réel du flux vidéo considéré. Au contraire, les algorithmes de tone mapping globaux utilisent une seule et unique fonction d'adaptation de luminance pour l'ensemble des pixels de l'image HDR, ce qui simplifie considérablement la complexité de leurs implantations matérielles et diminue de manière significative les ressources matérielles nécessaires. Pour cette raison, nous avons sélectionné l'algorithme de tone mapping global proposé par Duan et al. [Jiang Duan et al., 2010]. Ce choix est réalisé également au regard du nombre limité de paramètres devant être fixé par l'utilisateur (voir l'équation 5.1).

$$D_{ij} = C \times (D_{max} - D_{min}) + D_{max}$$

$$C = \frac{ln(E_{ij} + \tau) - ln(E_{ij(min)} + \tau)}{ln(E_{ij(max)} + \tau) - ln(E_{ij(min)} + \tau)}$$
(5.1)

Où  $D_{ij}$  est la luminance tone mappé de la luminance HDR  $E_{ij}$ .  $E_{ij(max)}$  et  $E_{ij(min)}$  sont respectivement le maximum et le minimum de luminance acquise de la scène.  $\tau$  est le paramètre de brillance (plus  $\tau$  est élevé plus l'image tone mappé est noir et vice-versa). Tableau 5.2 – Ressources utilisées par l'implantation matérielle de l'opérateur de tone mapping.

Type de composant	Nombre de composants	En %
Registres	3423	1%
LUT	2622	1%
Slice	842	2%
BUFG	1	3%

L'implantation matérielle utilise peu de ressources, moins de 1% de celles disponibles dans le FPGA considéré. Le module peut fonctionner à une fréquence maximale de 300*MHz*.

#### 5.2.4/ LA GÉNÉRATION HDR

La caméra HDR proposée utilise la technique de la fusion de plusieurs expositions à différents temps d'exposition dans le domaine de la radiance. Cette technique a été proposée par Debevec et Malik [Debevec et al., 2008]. Elle repose sur trois étapes : l'acquisition de plusieurs images LDR à différents temps d'exposition (elles sont capturées pour la même scène), l'obtention de la fonction de réponse de la caméra et enfin la fusion des données dans le domaine de la radiance en utilisant la fonction de réponse estimée. La méthode de Debevec et Malik [Debevec et al., 2008] est peu contraignante lors de la phase d'obtention de la fonction de réponse de la caméra. Cependant l'algorithme a démontré son efficacité en termes de la qualité de la dynamique imagée et sa simplicité permet d'envisager son implantation matérielle. Cette simplicité est due à la faible complexité des équations mathématiques mise en œuvre (équations (2.17,2.15,2.16).

Le logarithme de la fonction inverse de la caméra g() est utilisé pour convertir les

valeurs des pixels en valeurs de radiances équivalentes par l'utilisation de l'équation (5.2) obtenue de l'équation (2.15) en tenant en compte que le temps d'exposition  $\Delta t_j$  est connu. La fonction g() peut être utilisée pour déterminer les valeurs de la radiance, quelleque soit l'image acquise par le système d'acquisition associé à la fonction g().

$$lnE_{ij} = g(Z_{ij}) - ln\Delta t_j \tag{5.2}$$

Pour toutes les architectures matérielles qu'on proposera pour implanter le module de la génération HDR, les différentes valeurs des temps d'exposition et les différentes valeurs de la fonction inverse de la caméra sont tabulées dans les mémoires distribués du FPGA (mémoires LUTS - look up table) et représentées en format virgule flottante simple précision IEEE754. L'architecture de module HDR peut varier légèrement selon le module sélectionné pour réaliser la suppression d'artefact de type ghost, mais l'algorithme utilisé reste identique.

#### 5.3/ LE MODULE DE LA SUPPRESSION DE GHOST

Afin de supprimer l'artefact de ghost, nous avons implanté deux méthodes différentes : la méthode basée sur l'ordre des pixels et la méthode de pondération améliorée.

#### 5.3.1/ SUPPRESSION DE GHOST PAR LA MÉTHODE BASÉE SUR L'ORDRE DES PIXELS

Comme nous l'avons montré dans le chapitre précédent, la méthode basée sur l'ordre des pixels est la méthode la plus simple qui permet d'obtenir des résultats satisfaisants en termes de suppression de l'artefact de type ghost. L'implantation matérielle de la version sans opération de morphologie mathématique est la plus simple. Dans ce cas, une comparaison des deux flux LDR, basse et haute expositions, est réalisée avec le celui correspondant à l'exposition moyenne. Le résultat de cette comparaison permet de sélectionner, pixel par pixel, le flux le plus pertinent. L'architecture matérielle proposée pour ce module est illustrée dans figure 5.7.



FIGURE 5.7 – L'architecture du module de la suppression de l'artefact de ghost avec la méthode basée sur l'ordre des pixels.

Comme on peut observer sur la figure 5.7, le module ne fait que remplacer les flux de pixels LDR basse exposition et haute exposition respectivement  $Z_{ij}^1$  et  $Z_{ij}^3$  par le flux médian de moyenne exposition nommée  $Z_{ij}^{ref}$  et leurs temps d'exposition  $\Delta t_1$  et  $\Delta t_3$  par le temps d'exposition médian  $\Delta t_{ref}$  dans le cas où ces deux flux ne respectent pas leurs ordres par rapport au flux médian (si les valeurs des pixels de l'exposition faible sont supérieures à celles de l'exposition médiane , et inversement pour les valeurs de l'exposition forte). Ensuite, les flux vidéo sont envoyés directement au module de la génération HDR dont l'architecture matérielle est représentée dans la figure 5.8.

L'implantation matérielle de l'architecture proposée de système globale utilisant le module de suppression de l'artefact de ghost basé sur l'ordre des pixels, utilise 24% des ressources globales du FPGA. La fréquence maximale supportée par l'implantation matérielle est de  $117MH_z$ . Les ressources logiques limitées, au regard de solutions alternatives et la faible latence de génération HDR égale à 57 cycles d'horloge ( $0.495\mu s$  à



FIGURE 5.8 – Architecture de l'implantation matérielle de la génération HDR utilisant la méthode basée sur l'ordre des pixels utilisée pour la suppression de l'artefact de ghost.

114MHz) sont les avantages majeurs de cette architecture.

#### 5.3.1.1/ RÉSULTATS DE L'IMPLANTATION GLOBALE UTILISANT L'ORDRE DES PIXELS

Les résultats obtenus en utilisant notre caméra intégrant la suppression de ghost basée sur l'ordre des pixels, sont satisfaisants lorsque la scène filmée est bien éclairée et lorsque les temps d'exposition des différentes images LDR sont éloignés (voir la figure 5.9 à gauche). Cependant cette méthode donne des résultats impactés par le ghost et Tableau 5.3 – Ressources utilisées par l'implantation matérielle de l'architecture globale en utilisant le module de suppression de ghost selon l'ordre des pixels par rapport aux pixels de l'exposition moyenne qui est prise comme exposition de référence

Type de composant	Nombre de composants	En %
Registres	28254	9%
LUT	23699	15%
Slice	9185	24%
IO	210	35%
DSP48E1	20	2%
RAMB36E1	88	10%
RAMB18E1	15	1%
BUFG	12	37%



FIGURE 5.9 – Efficacité de la suppression de ghost en utilisant la méthode d'ordre des pixels, à droite dans des conditions de faible luminance et à gauche dans des conditions de forte luminance.

en même temps très bruités lorsque la scène filmée est faiblement éclairée.

#### 5.3.2/ SUPPRESSION DE GHOST PAR LA FONCTION DE PONDÉRATION

Afin d'implémenter la fonction de pondération permettant la suppression de l'artefact de ghost, nous l'avons simplifiée mathématiquement par l'équation 5.5. En effet, son implantation directe nécessite l'utilisation de plusieurs opérations exponentielles qui allongent le temps de traitement et complexifie l'implantation matérielle.

$$w(Z_{ij}^k) = \frac{[a(Z_{ij}^{ref})]^2}{[a(Z_{ij}^{ref})]^2 + [(\frac{f^{-1}(Z_{ij}^k)\Delta t_{ref}}{f^{-1}(Z_{ij}^{ref})\Delta t_k} - 1)]^2}$$
(5.3)

$$\frac{f^{-1}(Z_{ij}^l)}{\Delta t_l} = \exp\left[g(Z_{ij}^l) - \ln \Delta t_j^l\right]$$
(5.4)

$$w(Z_{ij}^k) = \frac{[a(Z_{ij}^{ref})]^2}{[a(Z_{ij}^{ref})]^2 + [\exp\left(g(Z_{ij}^k) - \ln\Delta_j^k - g(Z_{ij}^{ref}) + \ln\Delta_j^{ref}) - 1]^2}$$
(5.5)

La simplification proposée en équation 5.3 permet d'utiliser un seul opérateur exponentiel dans la fonction de pondération améliorée. Pour implanter l'opérateur exponentiel, on a choisi une approximation en exploitant les séries de Taylor. Ces séries permettent d'écrire l'exponentiel sous forme d'une somme représentant un polynôme d'ordre n (voir équation 5.7). Dans notre cas, la valeur transmise à cet opérateur est en effet bornée tel que décrit en équation 5.6, et par conséquent il est possible d'implanter l'opérateur de l'exponentiel en utilisant un nombre réduit de coefficients et le développement limité peut se faire au voisinage de 0 (en exploitant l'équation 5.8).

$$\max |(\ln E_i^k - \ln E_i^{ref} + \ln \Delta_j^k - \ln \Delta_j^{ref})| \le 3$$
(5.6)

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$
(5.7)

$$\exp(x) = 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \dots + \frac{1}{n!} + o(x^n)$$
(5.8)

Un développement limité en voisinage de zéro est cohérent puisqu'en effet les valeurs d'entrée de cet opérateur sont proches de zéro. Différentes approximations relatives à des ordres différents ont été considérées (voir la figure 5.10), afin de montrer l'ordre le plus fidèle à l'opérateur exponentiel dans la plage de nos valeurs en entrée. Comme illustré sur la figure 5.10, la qualité de l'approximation est excellente pour un ordre 12. En effet, l'erreur est de l'ordre de 0.016% pour les valeurs de radiance extrêmes |x| = 3 (ce sont les valeurs minimales et maximales de la radiance récupérée par notre caméra). L'erreur est d'ailleurs acceptable pour des valeurs en dehors de la plage de fonctionnement. Typiquement l'erreur est inférieure à 0,88% pour des valeurs au voisinage de |x| = 6. À partir d'un ordre 6, l'erreur devient plus importante en atteignant 3.35% pour la valeur |x| = 3.



FIGURE 5.10 – Le tracé des approximations de l'exponentiel en série de Taylor : en rouge l'exponentiel natif de MATLAB confondu avec, en bleu, le DL d ordre 12 et en vert le DL d'ordre 6 .

### 5.3.2.1/ IMPLANTATION MATÉRIELLE DE LA FONCTION DE PONDÉRATION UTILISANT LE DL DE L'EXPONENTIEL

Pour valider notre choix d'architecture et avant de procéder à une éventuelle implantation matérielle, nous avons implémenté l'approche choisie en utilisant MATLAB. Cette méthode de suppression de ghost est réalisée avant la génération des données à haute gamme dynamique en combinant la fonction de pondération de Debevec et Malik avec notre fonction de pondération modifiée. Par conséquent, il nous est apparu naturel de tester l'influence d'un tel changement sur le contenu HDR généré avant de tester son efficacité dans la suppression de l'artefact de ghost. En effet, les erreurs ajoutées par le développement limité de l'exponentiel peuvent compromettre le rendement HDR.

#### 5.3.2.2/ L'EFFICACITÉ DE LA GÉNÉRATION HDR

Afin de tester l'efficacité de la génération des données à haute gamme de dynamique, nous avons généré des images HDR en utilisant les trois méthodes : la méthode standard de Debevec et Malik, la méthode de Debevec et Malik combinée avec la fonction de pondération simplifiée en utilisant l'opérateur exponentiel de MATLAB et la dernière méthode en combinant la fonction de pondération de Debevec et Malik avec notre fonction de pondération modifiée implémentée en utilisant les développements limités (DL) d'ordre 12 de l'exponentiel.

La méthode utilisant les DL de l'exponentiel permet d'obtenir visuellement une bonne qualité d'image dans les zones fortement éclairées tout comme la version utilisant l'opérateur exponentiel natif de MATLAB (figure 5.11). Dans les zones faiblement éclairées, les trois méthodes ont des rendements visuels identiques.

Pour approfondir la comparaison des résultats, nous avons proposé une comparaison des contours obtenus en utilisant la fonction de pondération qui utilise l'exponentiel de MATLAB et ceux obtenus en utilisant la fonction de pondération utilisant l'exponentiel développe en séries de Taylor d'ordre 12. La version de notre méthode utilisant le DL de l'exponentiel (voir la figure 5.12 à droite ) permet d'obtenir des résultats très similaires à ceux résultant de la version utilisant l'exponentiel de MATLAB (voir la figure 5.12 à gauche) et ceci dans les deux régions fortement éclairée et sous-exposée.



FIGURE 5.11 – Les trois images LDR utilisées pour la génération HDR sont présentées sur la partie supérieure. La partie inférieure montre les images HDR calculées à partir des trois images LDR. De gauche à droite : le résultat de la méthode standard donné par Debevec et al., le résultat de la méthode d'adaptation de poids en utilisant l'opérateur exponentiel de MATLAB et le résultat de la même méthode d'adaptation de poids utilisant un développement limité de l'exponentiel d'ordre 12.



FIGURE 5.12 – Les contours détectés sur l'image obtenue par la méthode d'adaptation de poids utilisant l'exponentiel de MATLAB à gauche et ceux détectés sur l'image obtenue par la version utilisant le DL de l'exponentiel d'ordre 12 de notre méthode à droite.

#### 5.3.2.3/ L'EFFICACITÉ EN SUPPRESSION DE GHOST

Pour tester l'efficacité et l'impact de l'utilisation de DL de l'exponentiel dans la fonction de pondération permettant la suppression de ghost, nous avons utilisé une série d'images à faible gamme de dynamique contenant des objets en mouvement (figure 5.13, partie

supérieure). La fonction de pondération utilisant le DL de l'exponentiel d'ordre 12 et la version utilisant l'opérateur de l'exponentiel natif de MATLAB donnent toutes les deux de bons résultats (figure 5.13, partie inférieure) où elles suppriment efficacement l'artefact ghost où on ne voit aucun objet répété dans la scène (figure 5.13).



FIGURE 5.13 – Les images LDR (partie supérieure) nécessaires à la génération des images HDR (partie inférieure) en utilisant les méthodes de suppression de ghost par la fonction d'adaptation de poids en utilisant l'exponentiel de MATLAB et celle utilisant le DL de l'exponentiel respectivement de gauche à droite.

#### 5.3.2.4/ IMPLANTATION MATÉRIELLE DE LA FONCTION DE PONDÉRATION AMÉLIORÉE UTILISANT LE DL DE L'EXPONENTIEL

L'implantation matérielle est réalisée en exploitant le parallélisme implicite de l'algorithme afin de traiter les flux de pixels en temps-réel au minimum à la vitesse de capteur pour fournir les poids de ces pixels à chaque front montant de l'horloge du système. Ce résultat a été obtenu d'une part par l'intégration de deux modules matériels déterminant de manière simultanée la pondération respectivement pour les expositions à temps long et à temps court (la pondération de l'image LDR à temps d'exposition moyen est toujours égale à 1) et d'autre part par la mise en œuvre d'une architecture pipeline permettant le calcul de l'exponentielle et la tabulation des coefficients de DL (voir la figure 5.14). Un schéma simplifié d'une implantation matérielle de l'adaptation du poids est donné dans la figure 5.15.



FIGURE 5.14 – Un schéma simplifié de l'architecture matérielle du module d'exponentielle en DL d'ordre 12.

L'architecture proposée pour implanter l'exponentielle (figure 5.14) est générique, elle permet d'implanter un DL d'ordre n de l'exponentielle (dans la limite des ressources logiques disponibles au sein du FPGA). Si le module de l'exponentiel n'est utilisé que pour la pondération de la suppression de ghost, la première addition de 1 donnée dans la figure 5.14 peut être simplifiée avec la soustraction de -1 dans le module de calcul de poids de suppression de ghost. Ce qui permet de minimiser la latence de module de calcul de calcul du poids de pondération. La latence induite par ce module est ainsi réduite de 16 cycles d'horloge (le coût d'une soustraction et d'une addition en virgule flottante). En effet, le module réalisant les opérations de soustraction et d'additions en virgule flottante comporte une latence minimale de 8 cycles d'horloge. Cette simplification permet en outre de diminuer les ressources logiques utilisées.

Le tableau 5.4 montre l'utilisation des ressources logiques par l'implantation de l'exponentiel en DL d'ordre 12. Le coût en ressources logiques de ce module s'élève à 12% des ressources du FPGA sélectionné. Le coût élevé est lié à l'utilisation d'opérateurs arithmétiques (multiplication et addition ) en virgule flottante simple précision (voir le ta-

Type de composant	Nombre de composants	En %
Registres	19498	6%
LUT	18348	12%
Slice	5714	15%
DSP48E1	0	0%
RAMB36E1	0	0%
RAMB18E1	0	0%

Tableau 5.4 – Les ressources logiques utilisées pour développement de Taylor d'ordre 12

bleau 5.5). Le module de l'exponentiel peut fonctionner à une fréquence de 204MHz avec une latence de 104 cycles d'horloge. Ainsi dans notre configuration, cette latence correspond à un délai de 0.91  $\mu$ s pour la fréquence de fonctionnement fixée à 114MHz (soit la fréquence d'horloge globale de notre système).

Tableau 5.5 – Opérateurs arithmétiques utilisés par l'implantation matérielle de module de calcul d'exponentiel en DL d'ordre 12

Type d'opérateur	Nombre d'opérateurs	Latence (cycles d'horloge)
Addition	12	8
Multiplication	22	8

Tableau 5.6 – Les ressources logiques utilisées par le module de suppression de ghost utilisant le module exponentiel

Type de composant	Nombre de composants	En %
Registres	22673	7%
LUT	20958	13%
Slice	6396	16%
DSP48E1	0	0%
RAMB36E1	0	0%
RAMB18E1	1	1%

Le module de suppression de ghost consomme 26% des ressources logiques du FPGA. Tous les opérateurs arithmétiques utilisés par ce module (voir le tableau 5.7) sont à virgule flottante. Plus de détails sur les ressources utilisées sont donnés dans le tableau 5.6). Le module requiert en entrée les trois flux LDR pour pondérer les deux flux de la faible exposition et la plus haute exposition (le flux médian est utilisé comme flux de référence d'où son poids constant de 1). Le module peut fonctionner jusqu'à  $164MH_z$  ce qui reste plus grand que la fréquence de fonctionnement attendu pour tous les modules. La latence de ce module est de 164 cycles d'horloge équivalent  $1.44 \ \mu s$  à la fréquence 114*MHz* de système globale.



FIGURE 5.15 – Un schéma simplifié d'une implantation matérielle réalisant la fonction de la suppression de ghost par pondération.

Tableau 5.7 – Opérateurs arithmétiques utilisés par l'implantation matérielle de module de pondération utilisé pour la suppression de ghost

Type d'opérateur	Nombre d'opérateurs	Latence (cycles d'horloge)
Addition	14	8
Multiplication	24	8
soustraction	4	8
division	2	28

#### 5.3.2.5/ LE MODULE DE LA GÉNÉRATION HDR UTILISANT LE MODULE DE LA SUPPRES-SION DE GHOST

L'architecture proposée fusionne la radiance de trois images LDR pour en créer une seule HDR (voir la figure 5.16) en utilisant toujours l'algorithme de Debevec et Malik. L'implantation matérielle utilise le niveau de pipeline maximal permis dans les opérateurs arithmétiques flottants utilisés afin d'accélérer le processus de la génération.

Pour que la génération HDR, soit réalisée en temps réel et afin qu'elle puisse fournir un pixel HDR par cycle d'horloge, toutes les entrées (à part  $W_{dg}(Z_{ij}^1) W_{dg}(Z_{ij}^3)$ ) doivent transiter via des mémoires de type FIFO. d'une profondeur permettant de supporter la latence du module de suppression de ghost. La taille requise sera dans notre cas égale à 164. La largeur de bus des FIFOs sera relative à l'encodage. Cette utilisation de plusieurs



FIGURE 5.16 – L'architecture de l'implantation matérielle de la génération HDR reçoit en entrée les poids calculés par fonction de pondération de la suppression de l'artefact de ghost.

mémoires de temporisation augmente considérablement le coût en ressources logiques

nécessaires à l'implantation de ce module.

Tableau 5.8 – Opérateurs arithmétiques utilisés par l'implantation matérielle de module de la génération HDR sans le module de suppression de ghost utilisant l'exponentiel

Type d'opérateurs	Nombre d'opérateurs	Latence (cycles d'horloge)
Addition	4	7
Soustraction	3	7
Multiplication	5	7
Division	1	28
conversion virgule fixe/flottante	3	6

Tableau 5.9 – Les ressources logiques utilisées par le module de génération HDR incluant le module de suppression de ghost utilisant le module exponentiel

TYpe de composant	Nombre de composants	En %
Registres	53893	17%
LUT	50304	33%
Slice	16165	42%
DSP48E1	0	0%
RAMB36E1	6	1%
RAMB18E1	5	1%

L'implantation matérielle du module de la génération HDR dont l'architecture est illustrée dans la figure 5.16 est réalisée en intégrant le module de la suppression de ghost à l'intérieur de celui de la génération HDR. L'objectif de cette intégration est de minimiser l'utilisation des ressources logiques (principalement pour minimiser la taille des mémoires de temporisation des différents signaux). L'implantation requiert 33% des ressources globales disponibles sur le FPGA. Le module matériel peut fonctionner à 194*MHz* sans aucune violation d'intégrité temporelle des signaux, même si la contrainte sur l'horloge de fonctionnement est fixée à l'horloge de capteur d'image (114*MHz*).

L'architecture globale illustrée dans la figure 5.2 est implantée en matériel avec l'utilisation de module de suppression de ghost basé sur la méthode de pondération améliorée (exploitant le module d'exponentiel). Cette implantation a permis d'avoir les résultats montrés dans le tableau 5.10. Elle requiert 45% des ressources logiques programmables de circuit FPGA, et le système final peut fonctionner à la fréquence 139.899*MHz*.

Type de composant	Nombre de composants	En %
Registres	74987	24%
LUT	68479	45%
Slice	15705	55%
IO	210	35%
DSP48E1	20	2%
RAMB36E1	35	4%
RAMB18E1	15	1%
BUFG	10	31%

Tableau 5.10 – Ressources utilisées par l'implantation matérielle de l'architecture globale en utilisant le module exponentiel dans la fonction de suppression de ghost

#### 5.3.2.6/ RÉSULTATS D'ACQUISITION

La version de notre caméra utilisant le bitstream de l'implantation matérielle intégrant la suppression de ghost basée sur la pondération améliorée qui utilise à son tour le module de l'exponentiel (implémenté en développement limité d'ordre 12) a permis d'obtenir les résultats illustrés dans la figure 5.17.

Afin de comparer les résultats de notre caméra avec une génération HDR logicielle (avec et sans suppression de ghost), nous avons capturé trois images (faible, moyenne et haute exposition) d'une scène équivalente (on n'a pas utilisé le même flux LDR, car notre caméra ne permet pas d'envoyer le flux d'images HDR et le flux d'images LDR simultanément à l'ordinateur hôte) et nous avons généré à partir de la plate-forme logicielle DegEva les images HDR correspondantes (avec et sans suppression de ghost). Comme illustré en figure 5.17, à partir des images acquises via notre caméra, il est possible d'obtenir une dynamique importante (typiquement dans cette scène 108.9468 *dB*) tout en corrigeant les artefacts de type ghost. Plus précisément, on observe simultanément les détails des LEDs et l'intérieur des boîtes recevant très peu de lumière. Cependant, pour les objets en mouvement, notre caméra confirme les résultats obtenus lors de l'évaluation logicielle de la méthode implantée, elle supprime efficacement le ghost contrairement à une génération avec la méthode de Debevec native. En contrepartie, cette implantation requiert une augmentation importante des ressources matérielle en comparaison avec la génération d'un flux HDR sans correction des artefacts de ghost.



FIGURE 5.17 – Efficacité de la suppression de ghost en utilisant la méthode de pondération améliorée (intégrant le module exponentiel implanté en DL d'ordre 12) première ligne : 3 expositions différentes (0.63*ms*, 4.72*ms* et 15.72*ms*) de la scène capturée, deuxième ligne à gauche résultat obtenu par notre caméra, au milieu le résultat obtenu par le logiciel DegEva avec la même méthode pour une scène équivalente et à droite le résultat logiciel sans suppression de ghost est, troisième ligne des agrandissements sur les zones d'intérêt contenant de HDR et du mouvement.

#### 5.3.3/ SUPPRESSION DE GHOST PAR LA FONCTION DE PONDÉRATION TABULÉE

L'implantation matérielle (illustrée en figure 5.18) de la fonction de pondération de suppression de l'artefact de ghost  $w_{dg}(Z)$  décrite dans l'équation (4.3) est complètement tabulée, ce qui signifie que la réponse de la cette fonction est pré-calculée. Ce choix permet d'économiser des ressources matérielles, mais aussi de garantir d'une fréquence de fonctionnement importante en limitant l'utilisation d'opérateurs arithmétiques. Nous avons opté pour le choix de la tabulation, car l'implantation matérielle d'un opérateur exponentiel consomme beaucoup de ressources logiques (voir l'implantation précédente qui consomme 45% des ressources logiques disponibles sur le FPGA utilisé). De plus ce choix ne permet pas une fréquence de traitement élevée tout en augmentant la latence du système global. Les mémoires embarquées de type LUT sont utilisées pour sauvegarder les valeurs en virgules flottantes de logarithme de la fonction inverse de la caméra g(Z) (donnée en équation 2.15) et les valeurs en virgules flottantes de la fonction  $a(Z)^2$  qui sont calculées au préalable par un script MATLAB. Le calcul de la fonction de pondération nécessaire à la suppression des artefacts de ghost peut être ainsi reformulé. L'architecture matérielle correspondante est ainsi décrite figure 5.18. Le calcul est réalisé à partir des poids précalculés et stockés dans des tables LUTs. Les poids sont décrits sous un format flottant IEEE754 simple précision. L'architecture requiert notamment 7 diviseurs à virgule flottante, 4 additionneurs (soustracteurs) à virgule flottante et 2 multiplicateurs à virgule flottante. Ce module peut atteindre une fréquence de fonctionnement de 254.647MHz (soit plus de 150% d'accélération par rapport à la version utilisant l'implantation de l'exponentiel) (voir le tableau 5.11 pour plus de détails sur l'implantation matérielle de ce module).

Tableau 5.11 – Les ressources logiques utilisées par le module de suppression de ghost tabulé

Type de composant	Nombre de composants	En %
Registres	5444	1%
LUT	4317	2%
Slice	1414	3%
DSP48E1	0	0%
RAMB36E1	4	1%
RAMB18E1	3	1%

#### 5.3.3.1/ LE MODULE DE LA GÉNÉRATION HDR UTILISANT LE MODULE DE SUPPRESSION DE GHOST TABULÉ

L'implantation matérielle de module de la génération de la grande gamme de dynamique par la fusion des données est illustrée dans la figure 5.16. Elle est complètement tabulée et elle utilise une architecture avec un nombre maximal d'étages pour accélérer



FIGURE 5.18 – L'architecture matérielle de la fonction de pondération en utilisant la tabulation de l'exponentiel des valeurs d'entrées.

le processus de la génération. Les LUTs sont utilisées pour sauvegarder les différentes valeurs de la fonction g(Z) et le logarithme des temps d'expositions. Pour les temps d'expositions, la plage prévue s'étend entre 175.2*us* et 359*ms* avec un pas de 175.2*us* ce qui donne 2048 valeurs possible de temps d'exposition). Les détails sur l'utilisation des ressources logiques nécessaires à cette implantation matérielle sont fournis dans le tableau 5.12.

Tableau 5.12 – Les ressources logiques utilisées par l'implantation matérielle de module de la génération HDR (avec le module de la suppression de ghost tabulé) sur le FPGA utilisé (virtex6 (xc6vlx240t)

Type de composant	Nombre	En %
Registre	19292	6 %
LUT	16704	11%
Slice	5492	14%
DSP48E1	0	0%
RAMB36E1	14	1%
RAMB18E1	9	1%

Le module de la génération HDR intègre le module de la suppression de ghost. Les calculs dans ces deux modules requièrent des valeurs tabulées. Le module global ainsi développé requiert 11% des ressources logiques du FPGA. La fréquence maximale du circuit est de 235.668*MHz*. Le faible coût de l'implantation matérielle nous a mené à implémenter ce module sur différentes gammes de FPGA ( 5.13,5.14 ). Les implantations matérielles faites sur un FPGA possédant une large quantité de ressources matérielles cependant, des implantations sur des plateformes moins puissantes sont possibles. L'implantation matérielle utilisant des tabulations ne requiert qu'un nombre limité de ressources matérielles au regard des ressources disponibles sur des FPGAs puissants de la série 7 du fabricant Xilinx, tel qu'illustré dans le tableau 5.13. Ce module de générateur HDR corrigé peut être ainsi utilisé comme un simple accélérateur matériel et permettre l'intégration de post traitements au sein du même FPGA.

Tableau 5.13 – Les ressources logiques utilisées par l'implantation matérielle de module de la génération HDR (avec le module de la suppression de ghost tabulé) sur des FPGA performants et de grandes tailles

Le circuit FPGA	virtex7 (x	c7vx485t)	kintex7 (xc7k325t)	
La fréquence (MHz)	286.275		275 286.275	
Type de composant	Nombre	En %	Nombre	En %
Registre	19277	3%	19277	4%
LUT	16745	5%	16784	8%
Slice	6025	7%	5685	11%
DSP48E1	0	0%	0	0%
RAMB36E1	14	1%	14	1%
RAMB18E1	9	1%	9	1%
L'utilisation globale (%)	ļ	5	8	3

Cette gamme de FPGA dispose en effet d'un grand nombre de ressources logiques et les implantations peuvent fonctionner à plusieurs centaines de MHz.

Les résultats d'implantation de la génération HDR avec la suppression de ghost sur des FPGA aux ressources plus modestes (virtex 5, artix 7 et zynq 7020) que celui utilisé par notre caméra démontrent la possible utilisation de plateformes possédant des ressources Tableau 5.14 – Les ressources logiques utilisées par l'implantation matérielle de module de la génération HDR (avec le module de la suppression de ghost tabulé) sur des FPGA de petites tailles

Le circuit FPGA	virtex5 (x	c5vfx70t)	zynq7020	) (xc7z020)	Artix7(xc7	a100tcsg324-2)
La fréquence (MHz)	187.	310	236	6.868	2	03.102
Le nom de composant	Nombre	En %	Nombre	En %	Nombre	En %
Registre	19542	43%	19277	18%	19290	15%
LUT	16460	36%	16755	31%	16197	25%
Slice	6593	58%	5642	42%	5180	32%
DSP48E1	0	0%	0	0%	0	0%
RAMB36E1	NDF	NDF	14	5%	14	5%
RAMB18E1	NDF	NDF	9	3%	9	3%
L'utilisation globale (%)	1	1	;	36		31

logiques relativement faibles par exemple les artix 7 ou le zynq7020. En effet, la quantité de ressources logiques utilisées ne dépasse pas les 36 % (voir le tableau 5.14). Ceci laisse la possibilité d'intégrer toute l'architecture globale de notre caméra. Plus encore dans le cas de zynq 7020 qui est de la famille des FPGA APSoC qui ouvre plusieurs perspectives. Dans le cas de FPGA virtex 5 l'implantation utilise 11% identique que celui de l'implantation des deux modules sur le circuit FPGA virtex 6 de notre caméra, néanmoins la fréquence de fonctionnement est plus faible.

Tableau 5.15 – Ressources utilisées par l'implantation matérielle de l'architecture globale en utilisant le module exponentiel tabulé dans la fonction de suppression de ghost

Type de composant	Nombre de composants	En %
Registres	39550	13%
LUT	33402	22%
Slice	12437	33%
IO	210	35%
DSP48E1	20	2%
RAMB36E1	44	5%
RAMB18E1	19	2%
BUFG	10	28%

L'implantation matérielle de l'architecture globale en utilisant la technique de la tabulation dans la fonction de suppression de ghost requiert seulement 22% des ressources de FPGA. Le système peut fonctionner à une fréquence de 118.85*MHz*.


FIGURE 5.19 – Comparaison de coût en ressources logiques (registres, LUT, Slices, DSP, etc) des trois architectures matérielles du système global. L'architecture utilisant l'ordre des pixels est donnée en bleu, celle utilisant la fonction de pondération tabulée est donnée en rouge, et la courbe tracée en vert représente les ressources utilisées par l'architecture utilisant la fonction de pondération utilisant le développement en série de Taylor de l'opérateur exponentiel.

À titre indicatif, il est intéressant d'étudier l'évolution du coût en ressources logiques en fonction du choix d'implantation au niveau de la fonction de pondération. La figure 5.19 illustre cette évolution, et fait clairement apparaître une diminution significative des ressources matérielles requises pour l'utilisation des versions tabulées et celle d'ordre des pixels.

### 5.3.4/ RÉSULTATS OBTENUS PAR NOTRE CAMÉRA HDR UTILISANT L'ARCHITEC-TURE TABULÉE DE MODULE DE SUPPRESSION DE GHOST

Cette comparaison, bien que qualitative, permet d'illustrer les similitudes d'un point de vue perception visuelle des deux méthodes : tabulée ou non. Pour cela, nous avons capturé la même scène que celle d'évaluation de la version utilisant la fonction de pondération exploitant le module de l'exponentiel. Les résultats obtenus sont illustrés dans la figure 5.20.



FIGURE 5.20 – Comparaison des résultats de la caméra utilisant la version tabulée de la fonction de pondération à gauche et celle non tabulée à droite.

Les résultats affichés dans la figure 5.20, montre que les deux versions ont le même rendement visuel, et ceci est valable pour le rendement HDR et à la suppression de ghost. Le rendement identique est expliqué par le fait que l'architecture tabulée de la fonction de pondération préserve toute la précision des calculs (dans la limite des erreurs de calcul flottant simple précision) et l'architecture utilisant l'exponentiel implémenté en DL utilise un nombre de coefficients élevés (12) (DL d'un grand ordre) ce qui implique une précision élevée sur la plage des données concernée.

#### 5.3.4.1/ EFFICACITÉ EN GÉNÉRATION HDR

Pour tester l'efficacité de la génération HDR, nous avons mesuré le contraste local dans une zone à faible luminosité. Cette dernière permet de mesurer la capacité d'une caméra à restituer toute la dynamique d'une partie de la scène. Dans notre cas, elle est réalisée en calculant l'entropie d'une mire composée de 80 régions (ou patchs) possédant des niveaux de gris différents. L'image acquise de cette mire est illustrée en figure 5.21 pour deux éclairages différents d'un écart de 7*stops* ). Les niveaux de gris sont linéairement croissants avec un écart de 6.5*stops* entre le patch le plus sombre et le patch le plus clair. Le tracée de la variation de l'entropie selon les différents éclairages est donné dans la figure 5.22.



FIGURE 5.21 – Mire composée de 80 patchs de niveaux de gris linéairement croissants capturée par notre caméra HDR (à gauche à 0 stops et à droite à 7 stops).

Sur le tracé de l'entropie donné en figure 5.22, on voit que notre caméra a un rendement stable pour les zones de faible luminosité avec une légère perte dans les zones à forte luminosité (à partir de 8*stops*).



FIGURE 5.22 – La préservation de contraste de la caméra HDR dans les basses lumières.

Le contraste maximal obtenu avec notre caméra sur la figure 5.22 est de 7.27 bits . Cette

limitation est liée à l'utilisation des images tone mappée (elles sont codées sur 8*bits*) dans le calcul de l'entropie ce qui reste très bon vu que l'entropie maximale qu'on peut avoir sur une image LDR codées sur 8*bits* est 8*bits* à condition que dans l'histogramme de cette dernière ne manque aucune valeur des valeurs de pixel permises (de 0 à 255).

#### 5.3.4.2/ EFFICACITÉ EN PRÉSERVATION DE TEXTURE

Afin de tester la capacité de notre caméra à préserver les textures de la scène acquise, nous avons utilisé une scène constituée de deux motifs identiques riches en texture (texturés), l'éclairage d'un motif reste fixe pendant que celui de l'autre motif varie de 1*stops* par acquisition.

Les agrandissements au niveau des zones riches en texture sur les différentes images de la figure 5.23, montrent que la caméra propose un excellent rendu visuel (elle préserve une bonne partie des textures même après un écart de 7*stops* d'éclairage entre le motif 1 et 2). Néanmoins la préservation des textures diminue avec l'augmentation de l'écart d'éclairage entre les deux motifs. Cette diminution est due à l'opération de compression de dynamique exercée par l'opération de tone mapping utilisé, car il est global et ce type de tone mapping n'a pas de bon rendement en texture locale.

### 5.3.4.3/ EFFICACITÉ EN GÉNÉRATION HDR ET SUPPRESSION DE GHOST

Pour valider les résultats de notre caméra en intérieur, nous avons utilisé le dispositif HDR perceptuel en intérieur de chez DxOMark. Ce dispositif est une scène qui permet d'évaluer de façon perceptuelle la capacité d'une caméra (ou d'un capteur d'image) à restituer une scène possédant une grande gamme de dynamique. Les acquisitions ont lieu à l'intérieur d'une pièce dans un environnement entièrement contrôlé (éclairage, mouvement, etc.) et de façon reproductible. Elle est constituée de :

 a Une mire composée d'objets naturels, texturés et colorés, la mire est éclairée par un éclairage dont l'illuminant et l'intensité sont contrôlés.

- **b** Une image représentant une scène naturelle (le devant d'un immeuble, des passants, des arbres, du ciel avec des nuages) est éclairée par un DMX.
- c Deux objets en mouvement sont disposés dans la scène permettant de qualifier les artefacts de fantômes (ghost) qui peuvent apparaître lors de l'utilisation de la technique de la fusion de plusieurs images LDR pour générer les images HDR.

Nous avons testé la caméra avec cette scène en utilisant trois types d'éclairages et illuminant :

- Illuminant A (20 lux) : la gamme dynamique de la scène obtenue est de 13*stops* (78*db*)
- Illuminant TL84 (300 lux) : générant une gamme dynamique de la scène de 9stops (54db)
- **3.** Illuminant D65 (1000 lux) : la dynamique de la scène est de 7*stops* (42*db*)

Les images LDRs dans la figure 5.24, sont données à titre d'illustration d'une scène équivalente que celle filmée par notre caméra, elles sont capturées pour le premier éclairage générant 13*stops* de dynamique. Les images LDRs qui ont permis de générer les différentes images HDR affichées dans la figure 5.24 ne sont pas capturées, car notre caméra ne permet pas de capturer l'image HDR et les images LDR qui sont utilisées dans sa génération en même temps. La figure 5.24 montre qu'en rendement visuel, notre caméra récupère toute la gamme de dynamique contenue dans la scène à différentes illuminations où les détails de la scène sont préservés dans l'image HDR (on distingue facilement les différentes photos et les stylos et les autres objets de la scène). Et de l'autre côté, la scène simulant une vue sur l'extérieur est aussi récupérée avec tous ces détails où on voit toutes les parties clés de la scène (les nuages, le ciel, les personnes et les arbres y compris leurs feuilles). Le mouvement lui aussi est bien corrigé où on ne voit qu'une seule instance de l'objet en mouvement et ceci sans effet halo autour de l'objet en mouvement de rotation.

### 5.4/ CONCLUSION

Dans ce chapitre nous avons montré l'architecture matérielle de la caméra conçue pour la génération HDR sans artefact de mouvement. Les différentes implantations matérielles réalisées sont exhaustivement détaillées, l'utilisation des ressources matérielles de chacun des modules principaux de l'architecture globale présentée (et celle des sousmodules critiques comme l'implantation de module exponentiel)font l'objet de données chiffrées. De plus, nous avons donné une comparaison entre les différentes versions de l'architecture matérielle globale. Cette comparaison est réalisée selon deux approches, une matérielle via la comparaison de coût en ressources logiques et complexité architecturale et l'autre fonctionnelle selon le rendement en HDR et l'efficacité en suppression de l'artefact de mouvement (ghost).



FIGURE 5.23 – Scène de teste de rendu HDR pour l'évaluation de la préservation du contraste et de la texture pour des conditions d'éclairage de *0stops* à *7stops* (de la gauche haut à la droite bas) (scène réalisée par DxOMark).



FIGURE 5.24 – Les images LDR (partie supérieure) équivalentes aux images LDR utilisées pour la génération de l'image HDR (partie inférieure) en utilisant notre caméra (intégrant l'étape de la technique de tabulation dans la fonction de suppression de ghost) pour les 3 différents éclairages de la scène.

6

## **C**ONCLUSION ET PERSPECTIVES

À l'heure actuelle, le besoin de la grande gamme de dynamique accroît de plus en plus et dans plusieurs domaines d'application. Ceci est plus important dans le domaine de la vidéo surveillance et les systèmes de vision par ordinateur en général. Le moindre manque d'informations dans les images utilisées peut entraîner d'importants dégâts pouvant atteindre des vies humaines (dans le cas de la vidéo surveillance). L'étape de l'acquisition d'images HDR est primordiale pour de tels systèmes avec des contraintes élevées autant sur la quantité d'information (plus de détails dans l'image HDR) récupérée, mais aussi le nombre d'artefacts dans ces images. En plus de ces contraintes sur la qualité des images HDR, l'acquisition des images HDR doit être en temps réel ce qui impose d'autres contraintes sur le fonctionnement de système d'acquisition et ses performances.

La majorité des systèmes permettant la génération HDR existants sont logiciels, et ils ne répondent pas au critère du temps réel, ce qui rend leurs utilisations impossibles dans les domaines à fortes contraintes de temps réel. Des contraintes similaires empêchent aussi le peu de systèmes d'acquisition HDR qu'on peut trouver, car ces dernières souffrent d'un nombre d'artefacts élevé (pour les systèmes utilisant la technique de multiples expositions ou plusieurs capteurs d'image) ou d'un coût de fabrication élevé (pour les systèmes utilisant un capteur HDR dédié ).

La tâche qui m'a été confiée durant mes 3 années de thèse était la mise en œuvre

139

d'un système d'acquisition HDR (caméra HDR) qui intègre en plus de la génération HDR une étape de détection et de correction des artefacts liés aux mouvements dans la scène en temps réel. Afin de répondre aux différents et nombreux défis de la mise en place d'un tel système d'acquisition HDR, le travail de recherche présenté dans ce manuscrit est organisé dans l'ordre chronologique, où on a commencé par donner les notions de base de la génération HDR et les artefacts de mouvement qui l'accompagne suivies par les méthodes de détection et suppression de ces artefacts, puis l'étude des différentes méthodes de suppressions de ces artefacts en logicielle et on a fini par donner les détails sur la mise en place de la caméra HDR et les résultats obtenus en l'utilisant.

Le test et validation des méthodes de correction de la suppression des artefacts de mouvement, a été réalisé par le logiciel qu'on développé. Ce logiciel est un outil complet qui permet non seulement à évaluer les résultats des différentes méthodes de correction de ghost mais aussi toute la chaîne de la génération HDR et y compris l'affichage et la sauvegarde des résultats dans la majorité des formats d'image conventionnelle et même en format HDR. La méthode de pondération qu'on a modifiée et améliorée, a montré des bons résultats en suppression de l'artefact de mouvement par rapport aux autres méthodes étudiées. Trois versions de l'architecture globale ont été implantées en matérielle et détaillées dans ce manuscrit. Les trois versions de l'architecture globale fonctionnent en temps-réel et fournissent un flux vidéo de 60 images/secondes HDR ou tone mappées. La latence et le coût en ressources logiques occupées au sein de FPGA utilisé sont différents d'une version à l'autre et ceci selon la méthode de la suppression de ghost utilisée et l'approche d'implantation matérielle utilisée. La version utilisant la méthode de pondération tabulée est moins coûteuse en ressources logiques que la version utilisant la même méthode, mais sans tabulation (la même remarque pour la latence), la méthode utilisant l'ordre des pixels est la moins coûteuse des trois et la plus rapide. Les résultats de la version utilisant la méthode de pondération (tabulée et non tabulée) ont démontré l'efficacité de cette méthode en suppression de ghost et son adéquation à la cible matérielle.

### 6.1/ PERSPECTIVES

Les résultats obtenus par la caméra HDR qu'on a développée sont satisfaisants en qualité HDR et efficacité de suppression de l'artefact de ghost. Par contre, plusieurs améliorations sont envisageables. L'amélioration la plus pertinente est la correction de la couleur. Car cette dernière nécessite qu'elle soit pensée spécialement pour la HDR, car on a essayé des corrections avec les méthodes traditionnelles de calibration chromatique, mais ces dernières ont échoué, car elles ne sont pas compatibles avec le processus de la génération HDR en général et spécialement à partir de plusieurs images LDR. La seconde amélioration est l'utilisation d'une interface de communication externe plus rapide que le 1gbit/s Ethernet, car cette dernière ne peut pas supporter un flux vidéo HDR brut codé sur 32bit en plein résolution (1024 × 1280) à une cadence de 60 images/seconde. Donc l'utilisation d'une communication USB3.0 ou de 10Gbit Ethernet ou toute autre interface plus rapide que le 1Gbit Ethernet est nécessaire. Un capteur plus rapide est recommandé afin d'augmenter la cadence image. L'utilisation de ressources logiques au sein du FPGA est faible ce qui laisse la possibilité d'embarquer un traitement au sein du FPGA. Une nouvelle architecture a été pensée. Elle se base sur l'utilisation d'un FPGA SoC (Zynq Utlrascale). La nouvelle architecture tire bénéfice de l'architecture interne de ce système sur puce qui contient en plus de la partie FPGA (dite partie logique, la même que les FPGA traditionnels) un microprocesseur ARM quad cœur, un GPU et un codec H265 temps réel et d'autres composants sur une même puce. Cette architecture permet une répartition parfaite des différents modules et fonctions de l'architecture globale de notre

caméra HDR. Sur une telle architecture, les fonctions de contrôle et de communication seront assignées au microprocesseur ARM (piloté par des applications isolées ou sous un système d'exploitation embarqué). La création HDR et la suppression de ghost seront confiées à la partie logique de FPGA et le GPU se chargera d'éventuels post traitements et affichage. Enfin le codec H265 pour la compression de flux sortant (lorsqu' un flux vidéo compressé est demandé). Néanmoins cette nouvelle architecture n'a pas encore été validée expérimentalement, car la plate-forme ZCU102 ciblée ne permet pas d'utiliser le même module de la gestion de la mémoire vive pour permettre l'accès à deux images précédentes en même temps que l'acquisition courante. Des développements supplémentaires sont donc requis pour gérer efficacement la mémoire (à cause de la largeur du bus mémoire qui relie la partie logique de FPGA et la mémoire vive) et ainsi que les différentes interfaces de communications (DVI, SPI de capteur, Ethernet).

# GLOSSAIRE

- HDR : High Dynamic Range.
- HDRI : High Dynamic Range Imaging.
- LDR : Low Dynamic Range
- **CCD**: Charged Coupled Device.
- CMOS : Complementarity Metal-Oxyde-Semiconductor. Structure résultante de

l'association de transistors NMOS et PMOS.

- CIE : Commission Internationale de l'Éclairage
- SVH : Système de Vision Humaine
- CAN : Convertisseur Analogique Numérique
- CFA : Color Filter Array
- PCB : Printed Circuit Board
- ROI: Region Of Interest
- DSP: Digital Signal Processor
- CPU: Central Processing Unit
- GPU: Graphics Processing Unit
- SVD : Singular Value Decomposition
- RGB : Red Green Blue
- RGBE : Red Green Blue values with a one byte shared Exponent
- TIFF: Tagged Image File Format

- IEEE : Institute of Electrical and Electronics Engineers
- CRF: Camera Response Function
- PSNR : Peak Signal to Noise Ratio
- **MSE**: Mean Squared Error
- **UDP**: User Datagram Protocol
- TCP : Transmission Control Protocol
- FIFO : First In First Out
- APSoC : All Programable System on Chip
- SoC : System on Chip
- SNR : Signal to Noise Ratio. Rapport signal/bruit (RSB).
- SSIM : Structural Similarity Index Measurement.
- CI: Circuit Intégré.
- InGaAs : Indium Gallium Arsenic.

## BIBLIOGRAPHIE

- [Akahane et al., 2006a] Akahane, N., Ryuzaki, R., Adachi, S., Mizobuchi, K., et Sugawa, S. (2006a). A 200db dynamic range iris-less cmos image sensor with lateral overflow integration capacitor using hybrid voltage and current readout operation. Dans Solid-State Circuits Conference, 2006. ISSCC 2006. Digest of Technical Papers. IEEE International, pages 1161–1170. IEEE.
- [Akahane et al., 2006b] Akahane, N., Sugawa, S., Adachi, S., Mori, K., Ishiuchi, T., et Mizobuchi, K. (2006b). A sensitivity and linearity improvement of a 100-db dynamic range cmos image sensor using a lateral overflow integration capacitor. IEEE Journal of Solid-State Circuits, 41(4) :851–858.
- [Alachiotis et al., 2010] Alachiotis, N., Berger, S. A., et Stamatakis, A. (2010). Efficient pc-fpga communication over gigabit ethernet. Dans Computer and Information Technology (CIT), 2010 IEEE 10th International Conference on, pages 1727–1734. IEEE.
- [An et al., 2014] An, J., Ha, S. J., et Cho, N. I. (2014). Probabilistic motion pixel detection for the reduction of ghost artifacts in high dynamic range images from multiple exposures. EURASIP Journal on Image and Video Processing, 2014(1) :1– 15.

- [Banterle et al., 2009] Banterle, F., Debattista, K., Artusi, A., Pattanaik, S., Myszkowski, K., Ledda, P., et Chalmers, A. (2009). High dynamic range imaging and low dynamic range expansion for generating hdr content. Dans Computer graphics forum, volume 28, pages 2343–2367. Wiley Online Library.
- [Bergen et al., 1992] Bergen, J. R., Anandan, P., Hanna, K. J., et Hingorani, R. (1992). Hierarchical model-based motion estimation. Dans *European conference on computer vision*, pages 237–252. Springer.
- [Debevec et al., 1997] Debevec, P. E., et Malik, J. (1997). Recovering high dynamic range radiance maps from photographs. Dans *SIGGRAPH*, pages 369–378.
- [Debevec et al., 2008] Debevec, P. E., et Malik, J. (2008). Recovering high dynamic range radiance maps from photographs. Dans ACM SIGGRAPH 2008 classes, page 31. ACM.
- [El Gamal et al., 2005] El Gamal, A., et Eltoukhy, H. (2005). Cmos image sensors. *IEEE Circuits and Devices Magazine*, 21(3) :6–20.
- [F. Drago et al., 2003] F. Drago, K. Myszkowski, T. A., et Chiba, N. (2003). Adaptive logarithmic mapping for displaying high contrast scenes. EUROGRAPHICS 2003 / P. Brunet and D. Fellner, Volume 22 (2003), Number 3(3) :xx xx.
- [Fowler et al., 1998] Fowler, B., El Gamal, A., Yang, D., et Tian, H. (1998). A method for estimating quantum efficiency for cmos image sensors. Dans *Proc. SPIE*, volume 3301, pages 178–185.

- [Gallo et al., 2009] Gallo, O., Gelfand, N., Chen, W., Tico, M., et Pulli, K. (2009). Artifactfree high dynamic range imaging. Dans *Proceedings of the IEEE International Conference of Computational Photography (ICCP)*, San Francisco.
- [Gençtav et al., 2011] Gençtav, A., et Akyüz, A. O. (2011). Evaluation of radiometric camera response recovery methods. Dans SIGGRAPH Asia 2011 Posters, page 15. ACM.
- [Griffiths et al., 2017] Griffiths, D. J., et Wicks, A. (2017). High speed high dynamic range video. *IEEE Sensors Journal*, 17(8) :2472–2480.
- [Grosch, 2006] Grosch, T. (2006). Fast and robust high dynamic range image generation with camera and object movement. *Vision, Modeling and Visualization, RWTH Aachen*, pages 277–284.
- [Hanhart et al., 2015] Hanhart, P., Bernardo, M. V., Pereira, M., Pinheiro, A. M., et Ebrahimi, T. (2015). Benchmarking of objective quality metrics for hdr image quality assessment. EURASIP Journal on Image and Video Processing, 2015(1) :39.
- [Hara et al., 2005] Hara, K., Kubo, H., Kimura, M., Murao, F., et Komori, S. (2005). A linear-logarithmic cmos sensor with offset calibration using an injected charge signal. Dans Solid-State Circuits Conference, 2005. Digest of Technical Papers. ISSCC. 2005 IEEE International, pages 354–603. IEEE.
- [Harris et al., 1988] Harris, C., et Stephens, M. (1988). A combined corner and edge detector. Dans *Alvey vision conference*, volume 15, pages 10–5244. Manchester, UK.

- [Heo et al., 2010] Heo, Y. S., Lee, K. M., Lee, S. U., Moon, Y., et Cha, J. (2010). Ghostfree high dynamic range imaging. Dans *Asian Conference on Computer Vision*, pages 486–500. Springer.
- [Hosticka et al., 2003] Hosticka, B., Brockherde, W., Bussmann, A., Heimann, T., Jeremias, R., Kemna, A., Nitta, C., et Schrey, O. (2003). Cmos imaging for automotive applications. *IEEE Transactions on Electron Devices*, 50(1) :173–183.
- [Ingber, 2000] Ingber, L. (2000). Adaptive simulated annealing (asa) : Lessons learned. arXiv preprint cs/0001018.
- [Jacobs et al., 2008] Jacobs, K., Loscos, C., et Ward, G. (2008). Automatic highdynamic range image generation for dynamic scenes. *IEEE Computer Graphics and Applications*, 28(2).
- [Jiang Duan et al., 2010] Jiang Duan, MarcoBressan, C., et GuopingQiu (2010). Tonemapping high dynamic range images by novel histogram adjustment. *Pattern Recognition*, 43 :1847–1862.
- [Kalantari et al., 2013] Kalantari, N. K., Shechtman, E., Barnes, C., Darabi, S., Goldman,
  D. B., et Sen, P. (2013). Patch-based high dynamic range video. ACM Trans. Graph., 32(6) :202–1.
- [KAMAKURA, 2017] KAMAKURA, K. (2017). Image sensors meet leds. *IEICE Transactions on Communications*, 100(6) :917–925.
- [Kang et al., 2003] Kang, S. B., Uyttendaele, M., Winder, S., et Szeliski, R. (2003). High dynamic range video. Dans ACM Transactions on Graphics (TOG), volume 22, pages

319-325. ACM.

- [Kang et al., 2005] Kang, S. G., Woo, D. H., et Lee, H. C. (2005). Multiple integration method for a high signal-to-noise ratio readout integrated circuit. IEEE Transactions on Circuits and Systems II : Express Briefs, 52(9) :553–557.
- [Kronander et al., 2014] Kronander, J., Gustavson, S., Bonnet, G., Ynnerman, A., et Unger, J. (2014). A unified framework for multi-sensor hdr video reconstruction. *Signal Processing : Image Communication*, 29(2) :203–215.
- [Lapray et al., 2014] Lapray, P.-J., Heyrman, B., et Ginhac, D. (2014). Hdr-artist : An adaptive real-time smart camera for high dynamic range imaging. *Journal of Real-Time Image Processing*, pages 1–16.
- [Leñero-Bardallo et al., 2017] Leñero-Bardallo, J. A., Carmona-Galán, R., et Rodríguez-Vázquez, Á. (2017). A wide linear dynamic range image sensor based on asynchronous self-reset and tagging of saturation events. *IEEE Journal of Solid-State Circuits*.

[Litwiller, 2001] Litwiller, D. (2001). Ccd vs. cmos. Photonics Spectra, 35(1):154–158.

- [Lowe, 2004] Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. International journal of computer vision, 60(2) :91–110.
- [Lucas et al., 1981] Lucas, B. D., Kanade, T., et others (1981). An iterative image registration technique with an application to stereo vision.

- [Mangiat et al., 2010] Mangiat, S., et Gibson, J. (2010). High dynamic range video with ghost removal. Dans *Applications of Digital Image Processing XXXIII*, volume 7798, page 779812. International Society for Optics and Photonics.
- [Mangiat et al., 2011] Mangiat, S., et Gibson, J. (2011). Inexpensive high dynamic range video for large scale security and surveillance. Dans *Military Communications Conference, 2011 - MILCOM 2011*, pages 1772 –1777.
- [Mann et al., 1995] Mann, Picard, Mann, S., et Picard, R. W. (1995). On being 'undigital' with digital cameras : Extending dynamic range by combining differently exposed pictures. Dans *Proceedings of IS&T*, pages 442–448.
- [Mann et al., 2002] Mann, S., Manders, C., et Fung, J. (2002). Painting with looks : Photographic images from video using quantimetric processing. Dans *Proceedings* of the tenth ACM international conference on Multimedia, pages 117–126. ACM.
- [McIIrath, 2001] McIIrath, L. G. (2001). A low-power low-noise ultrawide-dynamicrange cmos imager with pixel-parallel a/d conversion. *IEEE Journal of Solid-State Circuits*, 36(5) :846–853.
- [Menzel et al., 2007] Menzel, N., et Guthe, M. (2007). Freehand hdr photography with motion compensation. Dans *VMV*, pages 127–134.
- [Mertens et al., 2009] Mertens, T., Kautz, J., et Van Reeth, F. (2009). Exposure fusion :
   A simple and practical alternative to high dynamic range photography. Dans
   *Computer Graphics Forum*, volume 28, pages 161–171. Wiley Online Library.

- [Miller et al., 2013] Miller, S., Nezamabadi, M., et Daly, S. (2013). Perceptual signal coding for more efficient usage of bit codes. *SMPTE Motion Imaging Journal*, 122(4) :52–59.
- [Min et al., 2009] Min, T.-H., Park, R.-H., et Chang, S. (2009). Histogram based ghost removal in high dynamic range images. Dans *Multimedia and Expo, 2009. ICME 2009. IEEE International Conference on*, pages 530–533.
- [Mitsunaga et al., 1999] Mitsunaga, T., et Nayar, S. (1999). Radiometric Self Calibration. Dans IEEE Conference on Computer Vision and Pattern Recognition (CVPR), volume 1, pages 374–380.
- [Miyagawa et al., 1995] Miyagawa, R., et Kanade, T. (1995). Integration-time based computational image sensors. *ITE Technical Report*, 19(65) :37–41.
- [Myszkowski et al., 2008] Myszkowski, K., Mantiuk, R., et Krawczyk, G. (2008). High dynamic range video, synthesis lectures in computer graphics and animation. *Mor*gan & Claypool Publishers.
- [Narwaria et al., 2015] Narwaria, M., Mantiuk, R., Da Silva, M. P., et Le Callet, P. (2015). Hdr-vdp-2.2 : a calibrated method for objective quality prediction of highdynamic range and standard images. *Journal of Electronic Imaging*, 24(1) :010501.
- [Nayar et al., 2003] Nayar, S. K., et Branzoi, V. (2003). Adaptive dynamic range imaging : Optical control of pixel exposures over space and time. Dans *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pages 1168–1175. IEEE.

- [ni et al., 2012] ni, Y., Bogdan Arion, Y. Z., et Potet, P. (2012). 120 db wdr cmos intensified camera for night vision. *OPTRO21012*.
- [Parzen, 1962] Parzen, E. (1962). On estimation of a probability density function and mode. The annals of mathematical statistics, 33(3) :1065–1076.
- [Pece et al., 2010] Pece, F., et Kautz, J. (2010). Bitmap movement detection : Hdr for dynamic scenes. Visual Media Production, Conference for, 0 :1–8.
- [Posch et al., 2011] Posch, C., Matolin, D., et Wohlgenannt, R. (2011). A qvga 143 db dynamic range frame-free pwm image sensor with lossless pixel-level video compression and time-domain cds. *IEEE Journal of Solid-State Circuits*, 46(1) :259–275.
- [Raman et al., 2007] Raman, S., et Chaudhuri, S. (2007). A matte-less, variational approach to automatic scene compositing. Dans *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pages 1–6. IEEE.
- [Reinhard et al., 2010] Reinhard, E., Heidrich, W., Debevec, P., Pattanaik, S., Ward, G., et Myszkowski, K. (2010). High dynamic range imaging : acquisition, display, and image-based lighting. Morgan Kaufmann.
- [Reinhard et al., 2002] Reinhard, E., Stark, M., Shirley, P., et Ferwerda, J. (2002). Photographic tone reproduction for digital images. ACM Transactions on Graphics, 21(3) :267–276.
- [Reinhard et al., 2005] Reinhard, E., Ward, G., Pattanaik, S., et Debevec, P. (2005). High Dynamic Range Imaging : Acquisition, Display, and Image-Based Lighting (The

Morgan Kaufmann Series in Computer Graphics). Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

- [Rhee et al., 2003] Rhee, J., et Joo, Y. (2003). Wide dynamic range cmos image sensor with pixel level adc. *Electronics Letters*, 39(4) :360–361.
- [Robertson et al., 2003] Robertson, M. A., Borman, S., et Stevenson, R. L. (2003).
  Estimation-theoretic approach to dynamic range enhancement using multiple
  exposures. Journal of Electronic Imaging, 12(2) :219–228.
- [Schrey et al., 2002] Schrey, O., Huppertz, J., Filimonovic, G., Bussmann, A., Brockherde,
  W., et Hosticka, B. J. (2002). A 1 k/spl times/1 k high dynamic range cmos image sensor with on-chip programmable region-of-interest readout. IEEE Journal of Solid-State Circuits, 37(7) :911–915.
- [Sen et al., 2012] Sen, P., Kalantari, N. K., Yaesoubi, M., Darabi, S., Goldman, D. B., et Shechtman, E. (2012). Robust patch-based hdr reconstruction of dynamic scenes. ACM Trans. Graph., 31(6) :203.
- [Shechtman et al., 2010] Shechtman, E., Rav-Acha, A., Irani, M., et Seitz, S. (2010). Regenerative morphing. Dans *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 615–622. IEEE.
- [Sidibe et al., 2009] Sidibe, D. D., Puech, W., et Strauss, O. (2009). Ghost Detection and Removal in High Dynamic Range Images. Dans *European Signal Processing Conference*, page 5, Glasgow, Royaume-Uni. 5 pages.

[Simakov et al., 2008] Simakov, D., Caspi, Y., Shechtman, E., et Irani, M. (2008). Summarizing visual data using bidirectional similarity. Dans *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE.

[Sony, 2008] Sony (2008). Ccd and cmos image sensors.

- [Srikantha et al., 2012a] Srikantha, A., et Sidibé, D. (2012a). Ghost detection and removal for high dynamic range images : Recent advances. *Signal Processing : Image Communication*, 27(6) :650 – 662.
- [Srikantha et al., 2012b] Srikantha, A., et Sidibé, D. (2012b). Ghost detection and removal for high dynamic range images : Recent advances. *Signal Processing : Image Communication*, 27(6) :650–662.
- [Storm et al., 2006] Storm, G., Henderson, R., Hurwitz, J., Renshaw, D., Findlater, K., et Purcell, M. (2006). Extended dynamic range from a combined linear-logarithmic cmos image sensor. *IEEE Journal of Solid-State Circuits*, 41(9) :2095–2106.
- [Sugawa et al., 2005] Sugawa, S., Akahane, N., Adachi, S., Mori, K., Ishiuchi, T., et Mizobuchi, K. (2005). A 100 db dynamic range cmos image sensor using a lateral overflow integration capacitor. Dans Solid-State Circuits Conference, 2005. Digest of Technical Papers. ISSCC. 2005 IEEE International, pages 352–603. IEEE.
- [Tocci et al., 2011] Tocci, M. D., Kiser, C., Tocci, N., et Sen, P. (2011). A versatile hdr video production system. ACM Transactions on Graphics (TOG), 30(4) :41.
- [Tom Mertens, 2007] Tom Mertens, Jan Kautz, F. V. R. (2007). Exposure fusion. Rapport technique, Hasselt University EDM transationale Universiteit Limburg Belgium, Univer-

sity College London UK.

- [Tursun et al., 2015] Tursun, O. T., Akyüz, A. O., Erdem, A., et Erdem, E. (2015). The state of the art in hdr deghosting : A survey and evaluation. Dans *Computer Graphics Forum*, volume 34, pages 683–707. Wiley Online Library.
- [Tursun et al., 2016] Tursun, O. T., Akyüz, A. O., Erdem, A., et Erdem, E. (2016). An objective deghosting quality metric for hdr images. Dans *Computer Graphics Forum*, volume 35, pages 139–152. Wiley Online Library.
- [Ward, 2003] Ward, G. (2003). Fast, robust image registration for compositing high dynamic range photographs from hand-held exposures. *journal of graphics, gpu, and game tools*, 8(2) :17–30.
- [Yadid-Pecht et al., 2003] Yadid-Pecht, O., et Belenky, A. (2003). In-pixel autoexposure cmos aps. *IEEE Journal of Solid-State Circuits*, 38(8) :1425–1428.
- [Yourganov et al., 2001] Yourganov, G., et Stuerzlinger, W. (2001). Acquiring high dynamic range video at video rates technical report dept. of computer science, york university may 2001.
- [Zhang et al., 2010] Zhang, W., et Cham, W.-K. (2010). Gradient-directed composition of multi-exposure images. Dans Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, pages 530–536. IEEE.

# TABLE DES FIGURES

2.1	Dessin illustrant l'angle solide, la courbe C dans l'espace forme l'angle	
	solide $\omega$ au point <i>P</i>	7
2.2	Radiance : une énergie incidente sur une surface unitaire dA à partir d'un	
	ensemble d'unités de directions $d\omega$ [Reinhard et al., 2010]	8
2.3	Le spectre de la lumière visible	9
2.4	Le tracé de la courbe d'efficacité lumineuse photopique donnée par le	
	CIE[Reinhard et al., 2010]	10
2.5	Fonctions de correspondance des couleurs pour les stimuli R, V et B de	
	CIE et l'observateur standard 2°[Myszkowski et al., 2008]	12
2.6	Fonctions de correspondance des couleurs pour les stimuli X, Y et Z de	
	CIE et l'observateur standard 2°[Myszkowski et al., 2008]	13
2.7	Le capteur d'image CCD et sa liaison au PCB de la caméra	16
2.8	Le capteur d'image CMOS et sa liaison au PCB de la caméra	16
2.9	Création d'image HDR à partir de multiples expositions LDR	24

2.10	La visualisation d'une seule scène avec plusieurs capteurs d'images via	
	un diviseur de faisceau[Griffiths et al., 2017]	24
2.11	La fonction de transfert d'un système d'acquisition d'image	25
2.12	Courbe de réponse inverse $g(Z)$ , pour la base de donnée d'images de	
	mémoriel de Debevec [Debevec et al., 2008]	26
2.13	Les fonctions de pondération normalisées associées aux courbes	
	de réponse citées [Mann et al., 1995, Debevec et al., 1997,	
	Mitsunaga et al., 1999, Reinhard et al., 2005].	29
3.1	Apparition de ghost suite au mouvement des objet dans les images LDRs	
	[An et al., 2014]	36
3.2	Apparition de ghost suite au mouvement de la caméra.	38
3.3	L'image de variance calculée pour trois images à différentes expositions.	
	[Jacobs et al., 2008]	42
3.4	Les cartes multi-seuils des deux images faiblement exposées (haut à	
	gauche) et hautement exposées (haut à droite) en prenant l'image de mi-	
	lieu comme référence [Min et al., 2009]	44
3.5	Le calcul de bitmap de 2 images statiques à différentes expositions donne	
	un même résultat[Pece et al., 2010].	46

3.6	La classification des différentes méthodes de détection et de sup-	
	pression de ghost selon les critères de classification donnés dans	
	[Srikantha et al., 2012b].	59
3.7	Le schéma de système HDR proposé par Nayar et Branzoi.	
	[Nayar et al., 2003]	59
3.8	La caméra AGD proposée par Nayar et Branzoi. [Nayar et al., 2003]	60
3.9	La caméra ADR proposé par Kronander et al. [Kronander et al., 2014]	61
4.1	L'interface graphique de logiciel GhostEva	69
4.2	L'image HDR tone mappée générée sans l'étape de suppression de ghost.	
	De haut en bas, dans la première image le ghost est généré par le mouve-	
	ment de cycliste et de la personne tout à gauche, Dans la deuxième image	
	le ghost est généré par le mouvement de tout les personnes présents dans	
	la scène. Dans la dernière image, le ghost est généré par le déplacement	
	de la boite et du stylo sur le bureau.	72
4.3	L'image HDR tone mappée générée en utilisant la méthode de déghosting	
	basée sur l'ordre des pixels avec une étape de dilatation sur la carte de	
	ghost	74
4.4	L'image HDR tone mappée générée en utilisant la méthode de déghosting	

basée sur l'ordre des pixels sans dilatation et sa carte de ghost détectée.

76

4.5	L'image HDR tone mappée générée en utilisant la méthode de déghosting	
	basée sur le MTB et une érosion et dilatation pour sa carte de ghost	77
4.6	L'image HDR tone mappée générée en utilisant la méthode de déghosting	
	basée sur le MTB sans opération de morphologie mathématique sur sa	
	carte de ghost.	79
4.7	L'image HDR tone mappée générée en utilisant la méthode de déghosting	
	basée sur la prédiction et sa carte de ghost.	81
4.8	L'image HDR tone mappée générée en utilisant la méthode de déghosting	
	basée sur la prédiction dans le domaine de radiance et sa carte de ghost.	83
4.9	L'image HDR tone mappée générée en utilisant la méthode de déghosting	
	basée sur la variance avec une étape de dilatation sur sa carte de ghost	84
4.10	L'image HDR tone mappée générée en utilisant la méthode de déghosting	
	basée sur la variance et sa carte de ghost.	86
4.11	L'image HDR tone mappée générée en utilisant la méthode de déghosting	
	basée sur la fonction de pondération simple originale	88
4.12	L'image HDR tone mappée générée en utilisant la méthode de suppression	
	de ghost basée sur la fonction de pondération simple modifiée	91
4.13	Graphique de la comparaison des résultats des différentes méthodes pour	
	la scène du bureau.	93

4.1	4 Graphique de la comparaison des résultats des différentes méthodes pour	
	la scène du bureau.	94
4.1	5 Graphique de la comparaison des résultats des différentes méthodes pour	
	la scène du bureau	95
4.1	6 Graphique de la comparaison des résultats moyens des différentes scène	
	pour toutes les méthode.	96
5.1	La caméra HDR	103
5.2	Hardware architecture of the described HDR camera	104
5.3	Le fonctionnement de l'unité de la gestion de la mémoire globale avec la	
	permutation de trois images LDR	105
5.4	Détails sur la gestion de flot de donnée réalisé par l'unité de gestion de	
	la mémoire MMU (les mémoires FIFOs sont représentées par des cadres	
	bleus et la DDR3 SDRAM par des cadres rouges )	106
5.5	Architecture du module de transmission des trames de données via l'inter-	
	face Ethernet	107
5.6	Trame de transmission des données d'une ligne	108
5.7	L'architecture du module de la suppression de l'artefact de ghost avec la	
	méthode basée sur l'ordre des pixels.	112

5.8	Architecture de l'implantation matérielle de la génération HDR utilisant la	
	méthode basée sur l'ordre des pixels utilisée pour la suppression de l'arte-	
	fact de ghost	113
5.9	Efficacité de la suppression de ghost en utilisant la méthode d'ordre des	
	pixels, à droite dans des conditions de faible luminance et à gauche dans	
	des conditions de forte luminance.	114
5.10	Le tracé des approximations de l'exponentiel en série de Taylor : en rouge	
	l'exponentiel natif de MATLAB confondu avec, en bleu, le DL d ordre 12 et	
	en vert le DL d'ordre 6.	116
5.11	Les trois images LDR utilisées pour la génération HDR sont présentées	
	sur la partie supérieure. La partie inférieure montre les images HDR cal-	
	culées à partir des trois images LDR. De gauche à droite : le résultat de	
	la méthode standard donné par Debevec et al., le résultat de la méthode	
	d'adaptation de poids en utilisant l'opérateur exponentiel de MATLAB et le	
	résultat de la même méthode d'adaptation de poids utilisant un développe-	
	ment limité de l'exponentiel d'ordre 12.	118
5.12	Les contours détectés sur l'image obtenue par la méthode d'adaptation	
	de poids utilisant l'exponentiel de MATLAB à gauche et ceux détectés sur	
	l'image obtenue par la version utilisant le DL de l'exponentiel d'ordre 12 de	
	notre méthode à droite.	118

	5.13 Les images LDR (partie supérieure) nécessaires à la génération des
	images HDR (partie inférieure) en utilisant les méthodes de suppression
	de ghost par la fonction d'adaptation de poids en utilisant l'exponentiel de
	MATLAB et celle utilisant le DL de l'exponentiel respectivement de gauche
119	à droite
	5.14 Un schéma simplifié de l'architecture matérielle du module d'exponentielle
120	en DL d'ordre 12.
	5.15 Un schéma simplifié d'une implantation matérielle réalisant la fonction de
122	la suppression de ghost par pondération.
	5.16 L'architecture de l'implantation matérielle de la génération HDR reçoit en
	entrée les poids calculés par fonction de pondération de la suppression de
123	l'artefact de ghost
	5.17 Efficacité de la suppression de ghost en utilisant la méthode de pondéra-
	tion améliorée (intégrant le module exponentiel implanté en DL d'ordre 12)
	première ligne : 3 expositions différentes (0.63ms, 4.72ms et 15.72ms) de
	la scène capturée, deuxième ligne à gauche résultat obtenu par notre ca-
	méra, au milieu le résultat obtenu par le logiciel DegEva avec la même mé-
	thode pour une scène équivalente et à droite le résultat logiciel sans sup-
	pression de ghost est, troisième ligne des agrandissements sur les zones
. 126	d'intérêt contenant de HDR et du mouvement.

- 5.19 Comparaison de coût en ressources logiques (registres, LUT, Slices, DSP, etc) des trois architectures matérielles du système global. L'architecture utilisant l'ordre des pixels est donnée en bleu, celle utilisant la fonction de pondération tabulée est donnée en rouge, et la courbe tracée en vert représente les ressources utilisées par l'architecture utilisant la fonction de pondération utilisant le développement en série de Taylor de l'opérateur exponentiel.
- 5.21 Mire composée de 80 patchs de niveaux de gris linéairement croissants
  capturée par notre caméra HDR (à gauche à 0stops et à droite à 7stops).
  133
- 5.22 La préservation de contraste de la caméra HDR dans les basses lumières. 133
- 5.23 Scène de teste de rendu HDR pour l'évaluation de la préservation du contraste et de la texture pour des conditions d'éclairage de 0*stops* à 7*stops*(de la gauche haut à la droite bas) (scène réalisée par DxOMark).

5.24	Les images LDR (partie supérieure) équivalentes aux images LDR utili-	
	sées pour la génération de l'image HDR (partie inférieure) en utilisant notre	
	caméra (intégrant l'étape de la technique de tabulation dans la fonction de	
	suppression de ghost) pour les 3 différents éclairages de la scène	138
A.1	Les images LDR de la scène de cycliste	175
A.2	Les images LDR de la scène de jardin.	176
A.3	Les images LDR de la scène de bureau	176
B.1	Le menu déroulant de méthode de génération HDR de logiciel GhostEva .	177
B.2	Le menu déroulant de méthode de tone-mapping de logiciel GhostEva	178
B.3	Le menu déroulant de méthode de deghosting de logiciel GhostEva	178
B.4	Le menu déroulant d'analyse d'images de logiciel GhostEva	179
B.5	Le menu déroulant d'affichage des résultat de logiciel GhostEva	180
B.6	Le menu déroulant d'enregistrement des résultat de logiciel GhostEva	180
C.1	Le module de la carte imprimé de capteur d'image	181
#### LISTE DES TABLES

2.1	Correspondance entre les grandeurs radiométriques et celle photométriques	11
2.2	La mesure des gammes dynamiques et leurs expression selon le contexte.	
	[Miller et al., 2013]	15
2.3	Les performances et caractéristiques clés des capteurs HDR proposés	
	dans l'état de l'art.	22
2.4	La comparaison de fiabilité entre les différentes méthode de récupération	
	de la fonction de réponse de la caméra [Gençtav et al., 2011].	28
2.5	Les formats HDR[Reinhard et al., 2010].	31
3.1	Les avantages et inconvénients des métriques objective de suppression	
	de ghost en imagerie HDR	65
4.1	Les valeurs des différentes métriques HDR obtenues pour les différentes	
	séquences sans l'étape de suppression de ghost	73
4.2	Les valeurs des différentes métriques HDR des résultats obtenus par la	
	méthode utilisant l'ordre des pixels pour la suppression de ghost en utili-	
	sant une dilatation avec un noyau de $3 \times 3$ pixels	74

4.3	Les valeurs des différentes métriques HDR des résultats obtenus par la	
	méthode utilisant l'ordre des pixels pour la suppression de ghost sans opé-	
	ration de morphologie mathématique	75
4.4	Les valeurs des différentes métriques HDR des résultats obtenus par la	
	méthode basée sur le MTB pour la suppression de ghost avec une érosion	
	(noyau $2 \times 2$ pixels) et une dilatation (noyau $3 \times 3$ pixels)	78
4.5	Les valeurs des différentes métriques HDR des résultats obtenus par la	
	méthode basée sur le MTB pour la suppression de ghost sans opération	
	de morphologie mathématique	79
4.6	Les valeurs des différentes métriques HDR des résultats obtenus par la	
	méthode basée sur la prédiction pour la suppression de ghost	80
4.7	Les valeurs des différentes métriques HDR des résultats obtenus par la	
	méthode basée sur la prédiction dans le domaine de la radiance pour la	
	suppression de ghost	83
4.8	Les valeurs des différentes métriques HDR des résultats obtenus par la	
	méthode basée sur la variance pour la suppression de ghost avec une	
	opération de dilatation (noyau de 3 × 3 pixels)	85
4.9	Les valeurs des différentes métriques HDR des résultats obtenus par la	

méthode basée sur la variance modifiée pour la suppression de ghost ... 86

4.10	) Les valeurs des différentes métriques HDR des résultats obtenus par la	
	méthode basée sur la fonction de pondération simple originale pour la sup-	
	pression de ghost	88
4.11	Les valeurs des différentes métriques HDR des résultats obtenus par la	
	méthode basée sur la fonction de pondération simple modifiée pour la sup-	
	pression de ghost. On présente ici non pas les cartes de ghost à droite (qui	
	n'existent pas pour cette méthode) mais les contributions des images LDRs	90
4.12	2 Tableau de comparaison des résultats des différentes méthodes pour la	
	scène du bureau	92
4.13	3 Tableau de comparaison des résultats des différentes méthode pour la	
	scène du jardin	94
4.14	1 Tableau de comparaison des résultats des différentes méthode pour la	
	scène de cycliste	95
4.15	5 Tableau de comparaison des résultats moyens des différentes scènes pour	
	toutes les méthodes	96
5.1	Ressources utilisées par l'implantation matérielle de module de communi-	
	cation éthernet	109
5.2	Ressources utilisées par l'implantation matérielle de l'opérateur de tone	
	mapping	110

5.3	Ressources utilisées par l'implantation matérielle de l'architecture globale	
	en utilisant le module de suppression de ghost selon l'ordre des pixels par	
	rapport aux pixels de l'exposition moyenne qui est prise comme exposition	
	de référence	114
5.4	Les ressources logiques utilisées pour développement de Taylor d'ordre 12	121
5.5	Opérateurs arithmétiques utilisés par l'implantation matérielle de module	
	de calcul d'exponentiel en DL d'ordre 12	121
5.6	Les ressources logiques utilisées par le module de suppression de ghost	
	utilisant le module exponentiel	121
5.7	Opérateurs arithmétiques utilisés par l'implantation matérielle de module	
	de pondération utilisé pour la suppression de ghost	122
5.8	Opérateurs arithmétiques utilisés par l'implantation matérielle de module	
	de la génération HDR sans le module de suppression de ghost utilisant	
	l'exponentiel	124
5.9	Les ressources logiques utilisées par le module de génération HDR in-	
	cluant le module de suppression de ghost utilisant le module exponentiel .	124
5.10	Ressources utilisées par l'implantation matérielle de l'architecture globale	
	en utilisant le module exponentiel dans la fonction de suppression de ghost	125
5.11	Les ressources logiques utilisées par le module de suppression de ghost	
	tabulé	127

5.12 Les ressources logiques utilisées par l'implantation matérielle de module	
de la génération HDR (avec le module de la suppression de ghost tabulé)	
sur le FPGA utilisé (virtex6 (xc6vlx240t)	128
5.13 Les ressources logiques utilisées par l'implantation matérielle de module	
de la génération HDR (avec le module de la suppression de ghost tabulé)	
sur des FPGA performants et de grandes tailles	129
5.14 Les ressources logiques utilisées par l'implantation matérielle de module	
de la génération HDR (avec le module de la suppression de ghost tabulé)	
sur des FPGA de petites tailles	130
5.15 Ressources utilisées par l'implantation matérielle de l'architecture globale	
en utilisant le module exponentiel tabulé dans la fonction de suppression	
de ghost	130

## ANNEXES

# A

## LA BASE DE DONNÉE D'IMAGES LDR UTILISÉE DANS LES TESTS LOGICIELS

Dans cette annexe, nous présenterons les bases de données d'images LDR utilisées pour évaluer les différentes méthodes de suppression de ghost. Les images LDR de test sont organisées de la plus basse exposition (à la gauche en haut) à la plus longue exposition (à la droite en bas) et ceci pour les trois différentes bases de donnée.



FIGURE A.1 – Les images LDR de la scène de cycliste [Srikantha et al., 2012a].



FIGURE A.2 – Les images LDR de la scène de jardin [An et al., 2014].



FIGURE A.3 – Les images LDR de la scène de bureau [Srikantha et al., 2012a].

B

### LES MENUS DÉROULANT DE LOGICIEL GHOSTEVA

Cette annexe présente les différents menus déroulant de logiciel développé et présenté

dans le chapitre 4.

HDRI Creating	
Select HDR:	
Debevec	
Exposure Fusion	
	▼ .
Lamda	No. of points
100	256
L	

FIGURE B.1 - Le menu déroulant de méthode de génération HDR de logiciel GhostEva

Tone-Mapping Operato	rs
Select operator :	
Reinhard TMO (Globa Reinhard TMO (Globa Reinhard TMO (Video Drago TMO Duan TMO Schlick TMO	I+Local)
TumblinRushmeier TM	0 🔻
0.18	8
Key value, a	Sharpening parameter, phi

FIGURE B.2 – Le menu déroulant de méthode de tone-mapping de logiciel GhostEva

None		
weighting_s		
Variance		
Prediction		
MTB local		
Prediction F		
Var Wfs		
-		
wstd_wfs		
wstd_wfs		
wstd_wfs		-
wstd_wfs	Erod_k	Ŧ
wstd_wfs Img_ref	Erod_k	•
Img_ref	Erod_k	•
wstd_wfs Img_ref	Erod_k	-
wstd_wfs Img_ref 3	Erod_k	-

FIGURE B.3 – Le menu déroulant de méthode de deghosting de logiciel GhostEva.



FIGURE B.4 – Le menu déroulant d'analyse d'images de logiciel GhostEva.



FIGURE B.5 – Le menu déroulant d'affichage des résultat de logiciel GhostEva.



FIGURE B.6 – Le menu déroulant d'enregistrement des résultat de logiciel GhostEva.

С

### LA CARTE IMPRIMÉ DE CAPTEUR D'IMAGE



FIGURE C.1 – Le module de la carte imprimé de capteur d'image.

#### **Résumé :**

La génération d'images à grande gamme de dynamique (HDR) à l'aide de plusieurs expositions est largement utilisée pour récupérer la gamme de dynamique entière d'une scène filmée. La technique se base sur la fusion de deux images (ou plus) à faibles gamme de dynamique(LDR). Cependant, cette technique doit être utilisée pour les scènes statiques et elle ne peut pas être appliquée directement pour les scènes contenant du mouvement. Les mouvements introduits par les objets dans les images de pile d'images LDR créent des artefacts fantômes dans l'image HDR reconstruite. Dans cette thèse, nous avons étudié et évalué un grand nombre d'algorithmes utilisés pour corriger ou éviter ces artefacts. Nous avons fait un compromis entre robustesse et complexité dans le choix de la méthode permettant la suppression de cet artefact afin de proposer un système de génération de vidéo HDR en temps réel (caméra intelligente). Ce dernier est implémenté sur un circuit FPGA. Cette caméra intelligente basée sur un FPGA est présentée avec des résultats expérimentaux de la conception démontrant l'efficacité de la méthode sélectionnée. Le système proposé permet de générer des flux vidéo HDR, y compris le processus de suppression des artefacts fantômes, à 60images/s pour une résolution de capteur complète ( $1280 \times 1024$ ).

Mots-clés : Multiples expositions, grande gamme de dynamique, systèmes temps-réel, détection d'artefacts de fantômes, algorithmes temps-réel, adaptation de poids, FPGA

#### **Abstract:**

()

High dynamic range (HDR) image generation using temporal exposure bracketing is widely used to recover the whole dynamic range of a filmed scene by fusion of two or more low dynamic range (LDR) images. Temporal exposure bracketing technique should be employed for static scenes and it cannot be applied directly for dynamic scenes. Motions introduced by moving objects in the LDR stack images create ghosts artifacts in the reconstructed HDR image. In this thesis, we have studied and evaluated a large nuber of algorithms used to correct or avoid these artifacts and we mad a trade-off between robustness and complexity in order to propose a real-time HDR video generation system. The real-time HDR image generation system is implemented on a FPGA circuit. This FPGA-based smart camera is presented with some experimental results to demonstrate the selected method and design efficiency. The proposed system enables HDR video streams, including ghost removal processing, to be generated at 60f ps for a full sensor resolution ( $1280 \times 1024$ ).

Keywords: Exposure bracketing, ghost detection, real-time algorithm, high dynamic range, smart camera, tone mapping, weight adaptation, FPGA

École doctorale SPIM - Université de Bourgogne/UFR ST BP 47870 F - 21078 Dijon cedex
 tél. +33 (0)3 80 39 59 10 ■ ed-spim@univ-fcomte.fr ■ www.ed-spim.univ-fcomte.fr

