2D Image Rendering for 3D Holoscopic Content using Disparity-Assisted Patch Blending

João Lino[1,2]*, Paulo Nunes*[1,3]*, Fernando Pereira*[1,2]

[1]Instituto de Telecomunicações, [2]Instituto Superior Técnico, [3]ISCTE - Instituto Universitário de Lisboa

E-mail: {joao.lino, paulo.nunes, fp}@lx.it.pt

Abstract

Holoscopy is a 3D technology that targets solving some of the limitations of current 3D technology, such as the need to wear special glasses to get the depth perception and the visual discomfort caused inherent accommodation issues. The display technologies available today at home are not yet compatible with 3D holoscopic image formats, for this reason methods that convert 3D holoscopic formats into 2D and current generation 3D image formats are required. There is, however, no fully automated, all-in-Focus and high performance 2D extraction method available in the literature. Moreover, the development of this type of conversion technology requires appropriate testing and assessment. Besides subjective testing with real people, there is nowadays no other reliable way to assess the performance of this type of conversion technology.

In this context, and after making an extensive review of 3D holoscopic capture and display technologies, this paper: a) proposes a novel, fully automated, 2D image extraction method for 3D holoscopic images and assesses its performance; and b) identifies potential No-Reference Image Quality Assessment metrics able to rate this type of 2D extractions and assesses their correlation with the human perception of quality.

The proposed 2D extraction method – *Disparity-Assisted Patch Blending* – outperforms, in “normal” conditions, all the available alternative methods. The NIQE metric seems promising as a potential candidate to reliably rate 2D extractions, although more research work should be done on BRISQUE as well to make it even more reliable.

Keywords - 3D Holoscopic, Disparity-Assisted Patch Blending, Micro-image patch size computation, Disparity estimation, No-Reference Image Quality Assessment

1 Introduction

For 3D video services to become practical and sustainable, adequate data formats for representing and delivering 3D video content considering different constraints are needed. Moreover, it is essential that factors minimizing the consumer quality of experience be avoided, such as viewing discomfort or fatigue and the need for wearing special gear. Holoscopy is a technology that solves the issues of previous and current 3D technology. The display technologies that exist today in homes are not yet compatible with 3D holoscopic image formats. To solve the issue of compatibility there is the need for methods that convert from 3D holoscopic formats into 2D or current generation 3D image formats.

The currently known conversion methods are either well defined and produce 2D images with artefacts, or partially defined promising to produce high quality 2D images. There is however no fully automated, well defined, 2D extraction methods described in literature.

An issue that arises from the development of this type of conversion technology is the testing of a proposed solution. Short of submitting it to tests with real people, there is no other documented way to know if a 2D extracted image has good or bad subjective quality. Part of the reason is that the image quality assessment metrics documented where never tested with 2D holoscopic extractions. The other part of the reason is because most documented and proven ways of assessing perceived quality in images require an original 2D image to compare the 2D extraction to. The problem is that in this conversion scenario there never is an original 2D image because the original is a 3D holoscopic image.

1.1 Objectives

This paper has the objective of presenting the following subjets:

- The basics of 2D extraction from 3D holoscopic images;
- A novel, fully automated, 2D image extractions method for 3D holoscopic images;
- Potential No-Reference Image Quality Assessment metrics able to rate 2D extractions;
- Performance of the proposed 2D extraction method;
- Correlation between the identified No-Reference Image Quality Assessment metrics and the human perception of quality.

2 Extracting Views from Holoscopic Imaging: a Review

Extracting 2D views from holoscopic images is a necessary process to appropriately display the captured light field information in 2D displays and in the various types of 3D (stereoscopic and auto-stereoscopic) displays nowadays available. In principle, displaying a full 3D holoscopic image [1] may be achieved by applying the inverse of the capturing process, this means by replacing the radiance sensor in a holoscopic camera by a flat panel display projecting the captured holoscopic image (see Figure 2.1). However, displaying part of the captured light field information in other types of displays requires processing the holoscopic image to extract the information in the appropriate format, notably: i) a single 2D image; ii) a stereo pair; and iii) multiple 2D views.

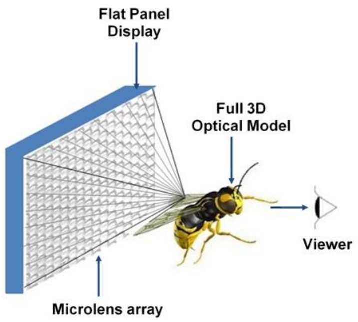
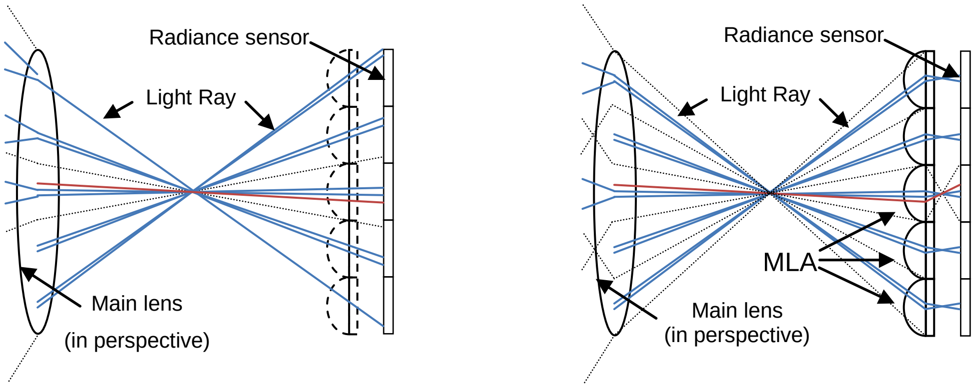


Figure 2.1 - Basic design of a light field display **[2]**.

Although the information required to reconstruct a 2D image of a scene is intrinsically included in the captured light field, the appropriate information must be selected and extracted to emulate the process of taking a 2D picture from a specific light field, in this case the captured light field. In traditional imaging, and holoscopic imaging alike, the radiance values captured represent a mathematical integral of light rays. To better understand the extraction problem addressed in this section, it is time to go deeper into this issue and better characterize the structure of the captured light field data, in order to have an idea how the extraction/reconstruction methods may work. In traditional imaging, the angular range of light rays allowed into the camera by the main lens is spread out through the radiance values captured by the radiance sensor, meaning that no angular coordinate ever repeats in the camera sensor. This phenomenon is illustrated by the ray trace diagram in Figure 2.2 a) where the traditional imaging case is represented. In holoscopic imaging, the same angular range repeats throughout all portions of the radiance sensor behind each micro-lens. This means that each micro-image consumes the total range of angles allowed into the camera by the main lens. Each of these combinations of angle coordinates is repeated as many times as the number of micro-lenses in the camera; Figure 2.2 b) presents this case.

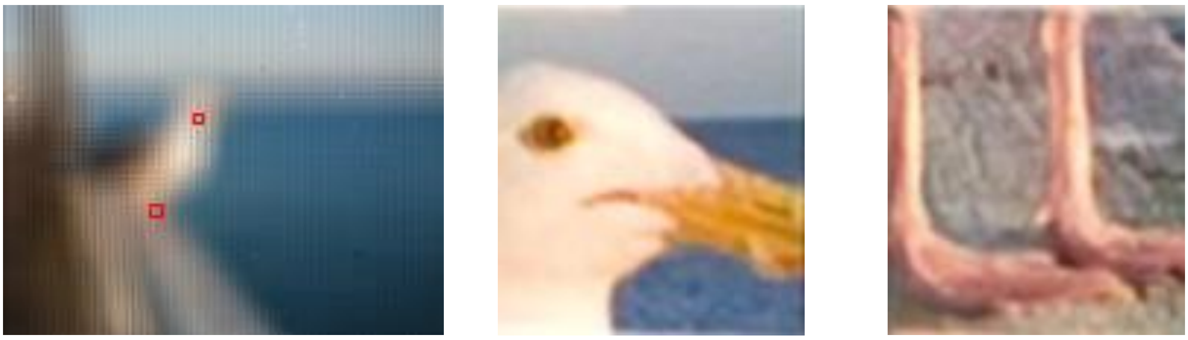


a) b)

Figure 2.2 - Light field capture: a) traditional imaging capture; b) holoscopic imaging capture. The Red, Blue and Gray ray traces in case a) and b) are the same cases in both images, only in b) there is a MLA in place. All rays come through the Main lens in the direction of the Radiance sensor. The Red ray traces represent rays that in case b) correspond to the most central PoV. The Gray ray traces represent the limit angle allowed into the camera through the Main lens, at its edge. The Blue ray traces represent, in case b), samples corresponding to far edge PoVs.

Because of these angular properties, it is appropriate to say that each micro-image corresponds to a single traditional image. To illustrate this point, from the holoscopic image in Figure 2.3 a) two micro-images were selected. The upper selection corresponds to the micro-image in Figure 2.3 b) and the lower selection corresponds to the micro-image in Figure 2.3 c). Although the angle ranges are the same for all micro-images, and equal to the total angular range the main lens allows, the PoV varies slightly among all micro-images.

The choice of data extracted from the captured holoscopic image to reconstruct a 2D image to represent the scene must be carefully made to obtain a representation as faithful as possible of the full scene because some regions of the scene repeat on adjacent micro-images.



a) b) c)

Figure 2.3 - Single micro-image representation: a) full holoscopic image; b) marked upper micro-image; c) marked lower micro-image.

3 The Proposed Disparity-Assisted Patch Blending 2D Extraction Algorithm

The proposed Disparity-Assisted Patch Bending 2D extraction (DAPBe) algorithm is presented in this section. The basic idea of this new algorithm is to extract 2D images from holoscopic images without the need to:

- Specify to the algorithm a particular plane of the scene that will be in focus in the extraction method. This algorithm generates images with the same depth of field of each micro-lens, typically a large depth of field, generating All-in-Focus 2D images;
- Manually improve the depth estimations calculated from the holoscopic image to improve the perceived quality of the final result. This algorithm is a fully automated solution to generate All-in-Focus 2D extractions from a holoscopic image.

The overall architecture of the image extraction method proposed in this section is presented in Figure 3.1.

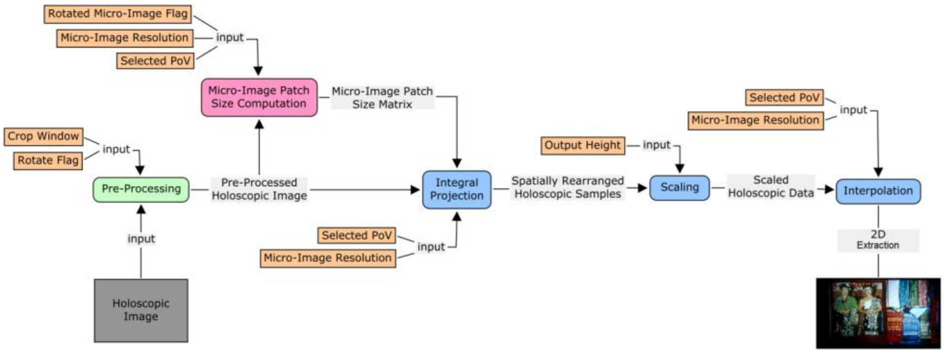


Figure 3.1 - Disparity-Assisted Patch Blending 2D image extraction architecture. Orange boxes represent inputs, while blue and pink boxes represent extraction modules with the pink boxes specifically related to disparity estimation.

The walkthrough of the proposed image extraction method is presented below. For each module, a brief description of the problem addressed as well as its major objectives and output(s) are also provided.

1. **Pre-Processing** – The problem this module tackles, e.g. micro-images rotation and misalignment, occurs due to the characteristics of some optical setups used in the acquisition process. For example:
    - Holoscopic images may exhibit a rotation of 180 degrees meaning that, depending on the relative position of the main lens focal plane relatively to the micro-lens array, the micro-images may appear inverted in the sensor plane;
    - Some micro-images may be incomplete at the borders of the holoscopic image due to some misalignments between the micro-lens array and the capturing sensor.

    Thus, the objective of this module is to align some of the most popular types of holoscopic data to be ready for processing with the proposed image extraction architecture since the proposed architecture expects complete non-rotated micro-images as input.

    The micro-images rotation problem is resolved by mirroring the samples over vertical and horizontal central lines. To ensure that only complete micro-images are considered in the following processing modules, a crop operation of the holoscopic image is performed around the image borders. The cropping is done manually according to the *crop window* input parameter. This value is one of the specifications of a holoscopic camera. It typically varies among different models and because there is no automatic means of detecting it, in this algorithm, it needs to be provided.

    The inputs and outputs of this module are the following:
    a. Input Holoscopic Image – Original holoscopic image;
    b. Input Cropping Window – Set of 4 values, each representing the distance in samples from the up, down, left and right limits of the holoscopic image, defining the cropping window;
    c. Input Rotate Flag – Binary flag indicating if the holoscopic image needs to be rotated or not;
    d. Output Pre-Processed Holoscopic Image – The cropped holoscopic image containing only complete micro-images rotated in order to have the scene oriented upwards.
2. **Micro-Image Patch Size Computation** – The problem this module tackles is the definition of micro-image patch sizes to guide the extraction process. Patch sizes are directly related to the extraction plane depth and directly related to the disparity of the central region of each micro-image regarding neighbouring micro-images. This disparity manifests horizontally in the left and right adjacent micro-images and vertically in the up and down adjacent micro-images.

    The objective of this module is to guide the Integral Projection module in the process of spatially arranging the holoscopic image samples according to their relative scene depth.

    To resolve the problem, the pre-processed image undergoes, first a sampling step - **Micro-Image Sampling** - at micro-image level, where central square portions, with different sizes, are gathered from each micro-image, followed by a disparity estimation step - **Disparity Estimation** - where a search process tries to find the best match in neighbouring micro-images for the central square portions sampled from each micro-image.

Micro-Image Sampling: The sampling is performed on the centre of each micro-image, where there is a representation of a *square portion A* of the scene, and on regions of the neighbours where it is very likely to find representations of that same region *A*; this happens when the selected PoV input points to the centre of the micro-images. If the selected PoV input points to another position (e.g., when the final objective is to extract a stereo pair, two different PoVs are defined, one for the left view and another for the right view) of the micro-images, then the sampling is performed in relation to that position and region representations in neighbouring micro-images.

Disparity Estimation: After the square portions have been sampled for each micro-image, a search process is performed to determine the disparity of the region representations within each micro-image, as it relates to their depth in the captured scene.

Using a Normalized Cross-Correlation criterion, for each micro-image, several square portions of the neighbours (the search region) are compared to the sampled *square portion A*, essentially searching for the *square portion A* in neighbouring micro-images. This disparity data if further processed to obtain a single value of disparity for each micro-image; thus disparity value expresses the spatial disparity, in samples, of the region represented in the centre of a micro-image and its neighbouring micro-images. The process of searching and subsequent processing of the data resulting from the search process, which results in the patch size value for each micro-image, will be described in more detail in the following section.

The inputs and outputs of this module are the following:

a. Input <u>Pre-Processed Holoscopic Image</u> – Output of the previous module which will be used as the source data for the sampling process;
b. Input <u>Micro-Image Resolution</u> – Two values, height and width, each is representing the vertical and horizontal dimensions in samples of the micro-images. This input is used to delimit the micro-images when gathering samples for the search process;
c. Input <u>Selected PoV</u> – A two dimensional coordinate pinpointing a position inside the micro-image and determining the region of the micro-images where the disparity estimation is performed; the selected PoV defines the centre of this region;
d. Input <u>Rotated Micro-Image Flag</u> – Binary flag responsible for informing the module if the image is rotated or not.
e. Output <u>Micro-Image Patch Size Matrix</u> – A patch size matrix defining the square patch size to be used for each micro-image in the extraction process.

3. **Integral Projection** – The problem this module tackles is the spatial arrangement of the 4D holoscopic samples into a simplified 2D image representation. Originally, the holoscopic samples have 4D spatial information, i.e., (y, z, φ, θ). In this module, the 4D spatial information is transformed into a 2D spatial representation by selecting the appropriate information as the output of this module.

   The objective of this module is to transform the 4D holoscopic samples into a “friendly” representation to the HVS and the display at hand, i.e., a 2D representation.

   To resolve the aforementioned problem, each micro-image is projected onto a 2D projection plane which can be abstracted as the plane where the 2D extracted image will be assembled from the available samples. The *Micro-Image Patch Size Matrix* input determines the size of a square region of samples centred on the Selected PoV of each micro-image that will correspond to the original size of a micro-image, on the projection plane, i.e. scaling up the micro-images through the projection process. The *Selected PoV* input determines the horizontal and vertical displacements of the projections, in the 2D projection plane, of each micro-image projection. The amount of displacement depends on the amount of scaling that was done in the projection, i.e., each micro-image is displaced proportionally to the distance of its projected samples; otherwise, all micro-image projections would be displaced by the same amount. This displacement can be seen as moving a real 2D camera horizontally or vertically, depending on the displacement direction, while still pointing it to a certain position in the scene. Because of the way the micro-images are scaled, they no longer occupy neatly defined positions in a matrix. Therefore, alternative means of storage need to be considered, which is covered in the next section.

The inputs and outputs of this module are the following:

a. Input Pre-Processed Holoscopic Image – A holoscopic image containing only complete micro-images, with the top of the scene oriented upwards; due to the structure of this input data, it can feed this module with holoscopic samples and also their position in the holoscopic image;
b. Input Micro-Image Patch Size Matrix – A patch size matrix defining the patch size to be used for each micro-image in the projection process.
c. Input Micro-Image Resolution – Two values, height and width, representing the micro-images dimensions; this input is used to delimit micro-images when asserting to which micro-image a sample belongs to;
d. Input Selected PoV – A two dimensional coordinate pinpointing a sample inside the micro-images. This sample results from the exposure to light from a particular angle (determined by camera optics). All these samples together (from all micro-images), makeup a scene representation from a particular PoV.
e. Output Spatially Rearranged Holoscopic Samples – Holoscopic samples and corresponding 2D coordinates on the 2D projection plane that forms a 2D representation of the captured scene, for the selected PoV. Because the samples may no longer have integer positions, they are no longer in matrix form but rather in an array form where each position has a sample, its new position in 2D space, and also the position of the sample inside the source micro-image, displaced by the *Selected PoV*.

4. **Scaling** – The problem this module tackles is to output a 2D image with a given spatial resolution out of the Rearranged Holoscopic Samples array produced by the previous module.

   The objective is to provide flexibility in the output resolution, while preserving scene proportions by respecting the ratio between horizontal and vertical sizes.

   To resolve the problem, the spatial information of the samples is scaled proportionally to the ratio of the *Pre-Processed Holoscopic Image*. Due to the ratio constraint, the final dimension of the output has to be set in relation to the width or height. Here, the height was chosen to facilitate compatibility with line resolution of a given display system; then the width is calculated by multiplying the ratio of the *Pre-Processed Holoscopic Image* with the *Output Height*.

   The inputs and outputs of this module are the following:

   a. Input Spatially Rearranged Holoscopic Samples – The output of the previous module.
   b. Input Output Height – The height of the output 2D reconstructed image; the width is determined by multiplying the ratio of the Pre-Processed Holoscopic image with this input value;
   c. Output Scaled Holoscopic Data – The scaled version of the *Spatially Rearranged Holoscopic Samples* input.

5. **Interpolation** - The problem this module tackles is to compute the 2D image samples for the positions of the output image matrix, the *2D Extraction*. In a "traditional" 2D image, image samples are arranged in a matrix fashion (regularly spaced). Although the samples composing the *Scaled Holoscopic Data* input are arranged in a 2D space, they are not regularly spaced, which means that some positions in the output image matrix are empty.

   The objective of addressing this issue is to create a full 2D image representation of the scene, i.e., a representation of the scene where samples are uniformly distributed in a matrix arrangement. To resolve the problem, an empty image is created to be populated by processing the information in the *Scaled Holoscopic Data* input. To that end, the input data is laid on top of the empty image to determine which micro-images lay on top of which empty sample positions. For each micro-image lying on top of a sample position, a sample is interpolated, from each micro-image, to that sample position. Further details on the interpolation method used in this algorithm will be covered in the next section. After all interpolated values are calculated, they are blended in a weighted average operation. The weight of each sample is determined through a weighting function following a Gaussian distribution. The samples closer to the displaced (by the Selected PoV input) centre of the corresponding micro-image have higher value than values far from the displaced centre of the corresponding micro-image. This is because the further from the displaced centre samples are, the less relevant to the selected PoV they are. Further details of this process will be covered in the next section.

The inputs and outputs of this module are the following:

a. Input Scaled Holoscopic Data – The output of the previous module, this means an incomplete 2D representation of the scene, where some sample positions are empty;
b. Output 2D Extraction – A complete 2D image representing the scene from a specific PoV with a specific spatial resolution.

4 Objective Evaluation

Automated IQA through objective evaluation is used not only in research but also in industry to guide image quality sensitive processes. To this end, it is important to have IQA metrics that are in line with human perception and rate images in a similar way as a human would. To date, no work has been done towards finding a suitable objective IQA metric that fits the task of evaluating 2D extractions from holoscopic images and rates them according to human quality perception.

Generally, IQA relies on a reference or “ground truth” image to provide a quantitative measure of comparison between an original and a ‘copy’ [3]. Because of this characteristic, methods that use a reference for the quality assessment are in a category known as Full Reference (FR) metrics. Popular quality metrics in this category include the *Peak Signal Noise Ratio* (PSNR), the *Root Mean Squared Error* (RMSE) [4] and *Structural SIMilarity* (SSIM) [5]. However, a reference image is not always available. When this is the case, other IQA methods exist to obtain the quality assessment: these methods are in a category known as No-Reference (NR) metrics.

Holoscopic images are not fit to use as a reference for any Full Reference metric that is available because 4D holoscopic images are very different from the 2D images that come out of the extraction methods. Thus, the only option to measure quality, resorting to objective IQA metrics, is to adopt NR IQA metrics.

Some NR IQA algorithms are based on models that can learn, through a training process, to predict human judgments of image quality from collections of human-rated ‘distorted’ images [6], [7], [8], [9], [10], [11]. Although they are in line with human perception, they are necessarily limited, since they can only assess quality degradations arising from the distortion types that they have been trained on. These algorithms are known as ‘opinion-aware’ (OA) because they have been trained on human rated distorted images and associated subjective opinion scores.

There are also NR IQA algorithms that do not base their score on human judgement. These algorithms are known as ‘opinion unaware’ (OU) and attempt to perform IQA based on distortion analysis metrics without human input. The OU algorithms in turn fall into two sub-categories depending on what type of distortion data is assumed or used in a training phase. Thus, an algorithm is classified as ‘distortion aware’ (DA) by design or by training on (and hence tuned to) specific distortion models to guide the QA process; algorithms classified as ‘distortion unaware’ (DU) rely only on exposure to naturalistic images or image models to guide the QA process.

Since no IQA metric has been tested or designed for the purpose of doing IQA on 2D extraction methods from holoscopic images, metrics from all categories of NR IQA metrics where considered as a possible source of objective IQA. The following list summarizes the motivations behind the choice of objective metrics, organized by the existing NR IQA categories, to be considered as 2D extraction IQA metric:

i. **Opinion Aware Distortion Aware** – Although many methods exist within this category, none is suited for this application because, in essence, this category represents 'built to purpose' NR IQA. Nearly all the available metrics target encoded content, which present very specific distortions that are not necessarily in line with the distortions found in 2D extractions. For this reason, no metric was considered appropriate and thus chosen from this category.

ii. **Opinion Aware Distortion Unaware** – This category may be a viable option for IQA within the holoscopic context because it does not rely on any specific distortion. Thus, algorithms like *Distortion Identification-based Image Verity and INtegrity Evaluation* (DIIVINE) [7], *Code Book Image Quality* (CBIQ) [9], *Learning Blind Image Quality* (LBIQ) [10], *BLind Image Integrity Notator using DCT-Statistics* (BLIINDS) [8] and *Blind/Referenceless Image Spatial QUality Evaluator* (BRISQUE) [6] are viable OA IQA solutions. As demonstrated in [8], from all these methods, BRISQUE is the one presenting the best overall performance [6] in comparison to other IQA metrics; for this reason, it was chosen.

iii. **Opinion Unaware Distortion Aware** – There is a single algorithm in this category that seemed relevant. The Anisotropic Quality Index (AQI) detects distortions associated with image integrity by calculating image anisotropy, and rates the content accordingly. Since the measured distortion relates to integrity, it may prove a good metric for NR IQA of 2D extractions. In testing for non-application specific IQA [12] [13], this method shows good performance against PSNR, SSIM and RSME; for this reason, it was chosen.

iv. **Opinion Unaware Distortion Unaware** – As far as NR goes, this is the category corresponding to the truly blind IQA algorithms. Of the few algorithms in this category, the *Natural Image Quality Evaluator* (NIQE) is the one worth mentioning at this time because it seems to hold up in testing [14] against other NR IQA metrics like BRISQUE.

From each category, an IQA metric was picked for testing in an attempt to find the best performing one for the 2D extraction problem at hand.

5 Comparing the 2D Extraction Methods Performance: Scores and Analysis

The VSBe and the SSPe typically perform below average and for this reason can't be considered the best methods (see Table 5.1). The DAPBe method is the only tested method able to produce *Excellent* quality 2D extractions, notably an extraction method that constructs a 2D representation of a scene with the same large depth of field of each micro-lens, typically called *All-In-Focus*. For its performance and algorithm of extraction, this author considers the DAPBe method the best in analysis. The DAPBe method is the one that extracts 2D image representation of a scene more faithfully to the original scene and therefore the best choice.

Table 5.1 - Percentage of MOS ratings attributed to each 2D extraction method

Method	Average and Above				Below Average		
	Excellent	Good	Fair	Total	Poor	Bad	Total
SSPBe	0%	18%	47%	65%	35%	0%	35%
DAPBe	20%	0%	40%	60%	40%	0%	40%
SSPe	0%	6%	18%	24%	29%	47%	76%
VSBe	0%	0%	0%	0%	14%	86%	100%

However, there are still subjects in the DAPBe method that need to be corrected for it to always perform better than the SSPBe method: a) the capacity to compensate for elliptical micro-images or not perfectly square micro-images; b) the capacity to compensate for errors in smooth regions of a scene.

In [15] are the 2D extractions of the available content captured by the Plenoptic 2.0 holoscopic camera. These All-in-Focus extractions serve as examples of the capability of the DAPBe method, by applying it to holoscopic images that have micro-images with a uniform structure.

6 Comparing the Objective Metrics Performance: Scores and Analysis

Based on the result obtained for the objective test scores and the apparent lack of connections between them and the subjective test scores further analysis is required to assess how much the objective and subjective test scores correlate with each other.

The Spearman's Rank Correlation coefficients and the Pearson Product-Moment Correlation Coefficients were calculated between the objective and the subjective ranks and scores respectively to determine if a hidden linear dependence exists. The scores obtained are presented in Table 6.1.

Table 6.1 – Spearman and Pearson correlation between the MOS and the objective test scores, for each extraction metric

Method	IQA	Spearman	Pearson
DAPBe	AQI	0,60	0,46
	NIQE	0,80	0,84
	BRISQUE	-0,20	-0,09
SSPBe	AQI	0,34	0,44
	NIQE	0,59	0,61
	BRISQUE	-0,27	-0,29
SSPe	AQI	-0,34	-0,13
	NIQE	0,43	0,62
	BRISQUE	-0,45	-0,29
VSBe	AQI	0,28	0,16
	NIQE	-0,37	0,36
	BRISQUE	-0,56	-0,45

A good result for the Pearson or Spearman correlations is typically above 0,9 or below -0,9, as can be noted by the results presented for all objective metrics considered, in [6], [7], [8], [9], [10], [12], [13] and [14]. Based on the scores presented in Table 6.1 that mark is never reached. In this regard there are grounds to conclude that all metrics have poor correlation with the MOS.

There is however enough data to choose the one that best correlates to the MOS ratings:

- For the DAPBe and SSPBe method - The NIQE rating is the best according to both Spearman and Pearson correlation;
- For the SSPe method – The BRISQUE rating is the best according to the Spearman correlation and the NIQE rating is the best according to the Pearson correlation;
- For the VSBe method - The BRISQUE rating is the best according to both Spearman and Pearson correlation;

Based on the fact that the NIQE ratings correlate best with MOS ratings within each group of single extraction method images, the NIQE metric is chosen as the Objective IQA metric that best correlates with the human perception of quality.

7 Conclusions

3D holoscopic technology has arrived at the consumer market and soon will hit the professional market as a consequence of ambitious projects like the 3D VIVANT Project. With the appearance of a company (and its products) called Lytro, everybody can now buy a 3D holoscopic camera and produce his own holoscopic content. However, because the display technology for 3D holoscopic images is not yet at an affordable price for the common buyer, Lytro had to limit its products to produce 2D images only. However, this should change in the future. For this change to happen, 3D holoscopy has not only to deliver a better experience than the previous 2D and 3D technologies, but it will also have to find some degree of compatibility with them. The currently available conversion methods to extract 2D images from 3D holoscopic images, solving part of the compatibility problem, still have performance issues. In this regard, there still is significant room for improvement.

Related to this issue, there is also the need to find quality metrics able to assess the perceivable quality of 2D extracted images, avoiding performing time consuming subjective quality tests with real people to score 2D extracted images. This kind of tool would play an important role on both 3D holoscopic technology development and deployment. In this context, this work has made the following contributions:

- **A novel, fully automated method for the extraction of 2D images from 3D holoscopic images** – The 2D extraction method proposed in this paper – *the Disparity-Assisted Patch Blending 2D extraction method* – is capable of both estimating the scene relative depth and, following that estimate, to extract a 2D representation of the scene based only on the specifications of the capturing camera and the 3D holoscopic capture itself. The process requires no human intervention during the reconstructions process and no input regarding any of the scenes depth properties. Moreover, the proposed 2D extraction solution outperforms, in "normal" conditions, all the available alternative methods; however, in "bad" conditions, it generates unwanted artefacts that compromise its performance making it similar to the other methods (in some cases, it may even be beaten by the SSPBe extraction method);
- **Potential No-Reference Image Quality Assessment objective metrics able to reliably rate 2D extractions from 3D holoscopic images** – While none of the no-reference objective image quality assessment metrics, the AQI, NIQE and BRISQUE, has been identified as a relevant potential candidate to reliably rate 2D extractions from 3D holoscopic images, NIQE seems to be promising. However, it is important to point out that both NIQE and BRISQUE need to be trained and are dependent on human opinion, although not in the same way. It may also be concluded that, with the originally trained model provided by its authors without considering 2D extractions, these NR IQA metrics can't reliably rate 2D extractions from 3D holoscopic images.

Bibliography

[1] A. Aggoun, "3D Holoscopic video content capture, manipulation and display technologies," in *2010 9th Euro-American Workshop on Information Optics*, 2010, pp. 1–3.

[2] A. Aggoun, E. Tsekleves, D. Zarpalas, A. Dimou, P. Daras, P. Nunes, and L. D. Soares, "Immersive 3D Holoscopic Video System," *IEEE Multimedia*, 2012.

[3] H. R. Sheikh, A. C. Bovik, and G. De Veciana, *An information fidelity criterion for image quality assessment using natural scene statistics.*, vol. 14, no. 12. Ieee, 2005, pp. 2117–2128.

[4] B. J. S. Armstrong and F. Collopy, "Error Measures For Generalizing About Forecasting Methods: Empirical Comparisons By J. Scott Armstrong and Fred Collopy Reprinted with permission form," *International Journal of Forecasting*, vol. 8, no. 1, pp. 69–80, 1992.

[5] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity.," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[6] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain.," *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society*, vol. 21, no. 12, pp. 4695–708, Dec. 2012.

[7] A. K. Moorthy and A. C. Bovik, *Blind Image Quality Assessment: From Natural Scene Statistics to Perceptual Quality*, vol. 20, no. 12. IEEE, 2011, pp. 3350–3364.

[8] M. Saad, A. C. Bovik, and C. Charrier, "Blind Image Quality Assessment: A Natural Scene Statistics Approach in the DCT Domain.," *IEEE Transactions on Image Processing*, vol. 21, no. 8, pp. 3339–3352, 2012.

[9] P. Ye and D. Doermann, "No-reference image quality assessment using visual codebooks.," *IEEE Transactions on Image Processing*, vol. 21, no. 7, pp. 3129–38, 2012.

[10] H. Tang and N. Joshi, "Learning a Blind Measure of Perceptual Image Quality," *Test*, vol. 1, pp. 305–312, 2011.

[11] T. Brandão and M. P. Queluz, *No-Reference Quality Assessment of H.264/AVC Encoded Video*, vol. 20, no. 11. 2010, pp. 1437–1447.

[12] S. Gabarda and G. Cristóbal, "Blind image quality assessment through anisotropy.," *Journal of the Optical Society of America A*, vol. 24, no. 12, pp. B42–B51, 2007.

[13] S. Gabarda and G. Cristóbal, "Image quality assessment through a logarithmic anisotropic measure," *SPIE Photonics Europe*, vol. 7000, no. 34, p. 70000J–70000J–11, Apr. 2008.

[14] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a 'Completely Blind' Image Quality Analyzer," *IEEE Signal Processing Letters*, vol. 20, no. 3, pp. 209–212, Mar. 2013.

[15] J. Lino, "Extractions of the Plenoptic 2.0 Holoscopic Camera Resources with the DAPBe Method," 2013. [Online]. Available: http://bordalo.img.lx.it.pt/apendix6.rar. [Accessed: 15-Oct-2013].