4D Light Field Disparity Map estimation using Krawtchouk Polynomials

Rui Lourenco*, Daniel Rivero-Castillo ‡, Lucas A. Thomaz*,†, Pedro A. A. Assuncao*,†,
Luis M.N. Tavora†, and Sergio M. M. de Faria *,†
*Instituto de Telecomunicações, Portugal
†ESTG - Instituto Politécnico de Leiria, Leiria, Portugal
‡Universidad Politécnica de Madrid, Spain
e-mails: {rui.lourenco, lucas.thomaz, amado, sergio.faria}@co.it.pt, daniel.rivero@upm.es, luis.tavora@ipleiria.pt

***Abstract*—This work presents an improved method to estimate disparity maps obtained from light field cameras using a novel edge detection algorithm based on Krawtchouk polynomials. The proposed method takes advantage of these polynomials to determine gradient information and find the edges based on automatically estimated weak and strong thresholds. The calculated edges in the gray scale epipolar plane image representation of a light field are then used to improve the accuracy of object boundaries in the the disparity map. The proposed method achieves better results when compared to other edge detection algorithms, both in terms of objective and subjective quality, specifically by reducing the mean squared error and the artifacts in the object boundaries. Furthermore, on average, the proposed method outperforms the state-of-the-art depth estimation algorithms, in terms of the objective quality of the final disparity map, namely for the commonly used HCI dataset.**

***Index Terms*—Light Field, Disparity, Depth, Edge Detection, Krawtchouk Polynomials, Structure Tensor**

I. Introduction

While conventional video cameras only capture light intensity, light field (LF) cameras also capture the direction of the light rays traveling in space towards the sensor. This allows an estimation of the depth information of the scene from a single image acquisition. Accurate depth information is important in a variety of contexts. Examples include medical diagnostic [1] and efficient rendering and coding of LFs [2].

Several algorithms have been proposed to extract such depth information from LF data. Tao *et al.* [3] estimate depth by mixing focus and correspondence cues. Lin *et al.* [4] extend this method with an iterative optimization step, while Williem *et al.* [5] enhance the focus and correspondence cues with a novel optimization step that takes into account the direction of scenic surfaces. Jeon *et al.* [6] use a sub-pixel extension of stereo correspondence to estimate disparity from multi-view images. Wanner *et al.* [7], propose the use of the structure tensor in order to estimate the orientation of quadrilateral structures in Epipolar Plane Images (EPIs), which is proportional to the disparity in LFs. This method proves to be remarkably accurate in homogeneous regions, while systematically failing in occlusion areas. This systematic error is designated silhouette enlargement [8], [9] as it results in an enlargement of the silhouettes of objects in the disparity map.

The quality of the edges determined by the structure tensor approach has been enhanced by the authors in [9], to improve the overall quality of the disparity maps. In that paper, a method is described that uses edge detection at the Epipolar Plane Image (EPI) level to detect the regions of the image with enlarged silhouettes and inpaint those regions with values from the background. However, since this is very reliant on the estimated edge maps (computed by the Canny algorithm [10]), visible artefacts appear near object boundaries whenever inaccurate edges are obtained.

This work further exploits the edge accuracy trend to improve the quality of the final disparity map. In this sense, a novel edge estimation algorithm is introduced for improving the accuracy of depth maps computed from 4D LFs. The proposed edge estimation algorithm uses Krawtchouk polynomials to calculate gradient information, which has been shown as capable of providing edges with higher accuracy than its counterparts [11]. In addition, the algorithm for inpainting erroneous regions is further refined. The overall results present a less segmented description of the object boundaries, i.e., less interrupted edges. Such improvements provide a clear reduction of artefacts, as well as in the mean square error, when compared to other non-supervised state-of-the-art methods.

The paper is organized as follows: in Section II, a background on LF images, epipolar plane images, disparity estimation and the Krawtchouk polynomials is given. In Section III, the proposed method for improving structure tensor-based disparity maps is presented. Section IV provides experimental results that compare the proposed disparity estimation method with other state-of-the-art algorithms. Finally, Section V draw some conclusions about this work.

This work was supported by Programa Operacional Regional do Centro, project PLenoISLA POCI-01-0145-FEDER-028325 and by FCT/MCTES through national funds and when applicable co-funded EU funds under the project UIDB/EEA/50008/2020, Portugal

978-1-7281-8750-1/20/$31.00 ©2020 IEEE

II. Background

This section aims to provide sufficient background knowledge on LF disparity estimation and discrete Krawtchouk orthogonal polynomials, as well as their potential for edge estimation that will be used in Section III in the proposed method. To this end, the section is divided in two subsections.

Subsection II-A provides a concise introduction to LFs, describing the epipolar plane image representation, as well as, a structure tensor based method for estimating disparity from LFs. Subsection II-B expands on the particular family of discrete polynomials called the Krawtchouk orthogonal polynomials.

A. Light Field, Epipolar Images and Disparity Estimation

A light field represents the flow of light on a scene with fixed illumination through unobstructed space. This can be conveyed by a 4-dimensional function $L(s,t,x,y)$, which can be seen as a 2D array of different views of the scene [12]. The spatial position of each viewpoint is represented by the index of the different views (s,t), while the pixel coordinates in each view are defined by (x,y).

Bolles *et al.* [13] observed that with enough density of viewpoints, x,s and t,y cuts of the LF, designated EPIs, contained relevant geometric information about the 3D scene, showing that the slope of a line in the EPI is inversely proportional to the depth of the point represented by that line.

As an edge in an EPI represents a single point of the scene across different view-points, estimating edges from these images proves more accurate and less prone to inaccuracies when compared to estimating edges directly in the view images. This is also true when estimating edges in the disparity map. Therefore, it is beneficial to use a method for depth map estimation that provides a full 4D depth map without a great deal of complexity increase. One of such methods is the structure tensor [14].

First proposed by J. Bigun in [15], the structure tensor is a matrix whose eigenvectors approximate an average direction of all lines in an image. When applied to an EPI $I^{(y)}(x,s)$, the structure tensor for each point (x,s) is given by:

$$\mathbf{T}(x,s) = \begin{bmatrix} I_x^2 * G_\sigma & I_x I_s * G_\sigma \\ I_x I_s * G_\sigma & I_s^2 * G_\sigma \end{bmatrix} = \begin{bmatrix} J_{xx} & J_{xs} \\ J_{xs} & J_{ss} \end{bmatrix}, \quad (1)$$

where $I_x = \frac{\partial I}{\partial x}$ and $I_s = \frac{\partial I}{\partial s}$ and G_σ represents a Gaussian kernel with scale σ. As the direction of linear symmetry corresponds to the orientation of EPI structures, one of the eigenvectors $\mathbf{e} = (e_0, e_1)$ is approximately co-linear with the lines located in the vicinity of the chosen point. Thus, the disparity of the chosen point can be calculated as $d = \frac{e_1}{e_0}$.

Furthermore, Bigun also describes a coherence measure $(r^{\cdot,\cdot}(\cdot,\cdot))$ for the structure tensor. The coherence, is a value bounded between zero and one that assesses how much the structures in the vicinity of a pixel of the image have a well defined direction. This is used as a reliability measure, which is useful for the optimization step described in Subsection III-D.

As discussed in [9], when the vicinity of any given point includes an occlusion edge belonging to a different object, some problems arise. In such cases, this point is consistently assigned the disparity of the neighboring occluding object. The consistency of these errors provides a good opportunity for improvement, which is explored in Section III.

B. Krawtchouk Polynomials

The basic theory on Krawtchouk polynomials is established in [16], [17], for the continuous version and in [18], [19], for the discrete version. Rivero-Castillo *et al.* proposed to use the discrete Krawtchouk Polynomials for edge estimation in [11].

Let there be a normalized weight function $\mu(x),\ x \in \Lambda$ such that $\Lambda = \{x_0, x_1, \ldots, x_N\} \subset \mathbb{R}$ with $x_0 < x_1 < \ldots < x_N$ and

$$\sum_{k=0}^{N} \mu(x_k) = 1. \quad (2)$$

The inner product of two functions f and g associated with the set Λ and weight function μ is defined as:

$$\langle f, g \rangle_{\Lambda,\mu} = \sum_{k=0}^{N} f(x_k) g(x_k) \mu(x_k) \quad (3)$$

In this manner, the monic Krawtchouk polynomials $k_n^\alpha(x,N)$ are orthogonal polynomials with respect to the binomial distribution.

$$\langle k_n^\alpha(x,N), x^j \rangle_{\Lambda_N, w_{N,\alpha}} = \sum_{i=0}^{N} k_n^\alpha(i,N) x^j w_{N,\alpha}(i) = 0, \quad (4)$$

for all $j = 0, 1, \ldots, N$, with

$$w_{N,\alpha}(x) = \binom{N}{x} \alpha^x (1-\alpha)^{N-x}, \quad (5)$$

where $x \in \Lambda_N$, with $\Lambda_N = \{0, 1, \ldots, N\}$, $N \in \mathbb{N}$, and $\alpha \in (0,1)$.

Let there be a function f of one variable. The first order difference of f can be defined as $\Delta_+ f(x) = f(x+1) - f(x)$, and the central difference between two nodes is given by the expression $\Delta f(x) = \frac{1}{2}\left(f(x+1) - f(x-1)\right)$. Krawtchouk polynomials allow closed expressions for the discrete derivatives of the polynomials such that:

$$\begin{aligned} \Delta_+ k_n^\alpha(x,N) &= n k_{n-1}^\alpha(x, N-1), \quad (6) \\ \Delta k_n^\alpha(x,N) &= \frac{n}{2}\left(k_{n-1}^\alpha(x,N-1) + k_{n-1}^\alpha(x-1,N-1)\right). \end{aligned}$$

These difference formulas can be extended to the two dimensional case $\mathbf{K}_{n,m}^{\alpha_1,\alpha 2}(x,y) = k_n^{\alpha_1}(x,N) k_n^{\alpha_2}(y,N)$, obtaining:

$$\begin{aligned} \Delta_x \mathbf{K}_{n,m}^{\alpha_1,\alpha 2}(x,y) &= \left(\Delta k_n^{\alpha_1}(x,N)\right) k_m^{\alpha_2}(y,N) \\ \Delta_y \mathbf{K}_{n,m}^{\alpha_1,\alpha 2}(x,y) &= k_n^{\alpha_1}(x,N) \left(\Delta k_m^{\alpha_2}(y,N)\right), \end{aligned} \quad (7)$$

where $\alpha_1, \alpha_2 \in (0,1)$, and $(x,y) \in \Lambda_N \times \Lambda_N$, [20], [21]. In Subsection III-A, these difference formulas are used to aid in the estimation of edges in an epipolar plane image.

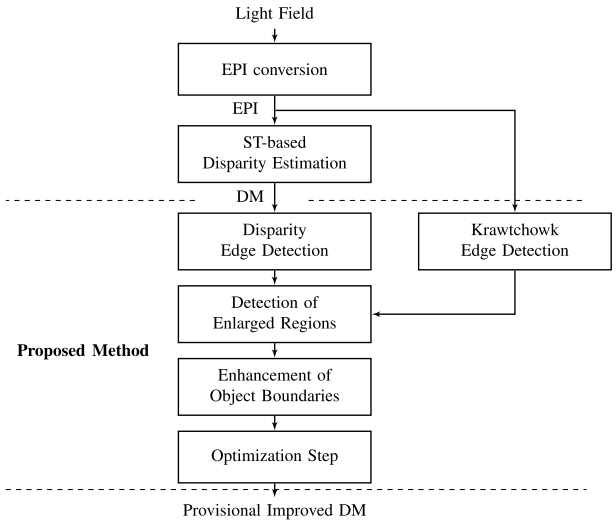

Fig. 1. Algorithmic structure of the proposed method.

III. Proposed Method

The proposed framework minimizes the silhouette enlargement effect by comparing edge maps estimated from the LF's EPIs with edge maps obtained from their corresponding disparity maps. As can be seen in the algorithmic structure represented in Figure 1, the proposed method receives as inputs horizontal and vertical EPIs of the form $I^{(y)}(x,s)$ and $I^{(x)}(y,t)$ and their respective structure tensor-based (ST-based) disparity maps (DMs) $d^{(y)}(x,s)$ and $d^{(x)}(y,t)$. Additionally, a reliability measure for the disparity maps $r^{(y)}(x,s)$ and $r^{(x)}(t,y)$ is provided.

The **Krawtchouk Edge Detection** block (Subsection III-A) provides an estimated edge map for a given EPI. The **Disparity Edge Detection** block (Subsection III-B) provides edge detection on the given disparity maps, these are used in the **Detection of Enlarged Regions** block (also Subsection III-B) while the **Enhancement of Object Boundaries** block (Subsection III-C) uses both edge estimations to enhance object boundaries in the disparity map. Finally, the **Optimization Step** block (Subsection III-D) describes a matting-based optimization step.

A. Krawtchouk Polynomial-based Edge Detection

Since the disparity maps input to the proposed algorithm (Fig. 1) present inaccurate object boundaries, it is necessary to correct them in order to obtain good estimations of the true location of object boundaries. In this paper, the use of an edge detection method based on Krawtchouk Polynomials on the EPIs is proposed to this end.

In order to approximate gradient information using Krawtchouk polynomials, a least squares approximation of the EPI can be computed as:

$$\mathbf{P}_{M_1,M_2}(x,s) = \sum_{n=0}^{M_1-1}\sum_{m=0}^{M_2-1} \beta_{n,m}\mathbf{K}_{n,m}^{\alpha_1,\alpha 2}(x,s), \qquad (8)$$

with

$$\beta_{n,m} = \frac{\langle I^{(y)}(x,s), \mathbf{K}_{n,m}^{\alpha_1,\alpha 2}\rangle_{2D}}{\langle \mathbf{K}_{n,m}^{\alpha_1,\alpha 2}, \mathbf{K}_{n,m}^{\alpha_1,\alpha 2}\rangle_{2D}}. \qquad (9)$$

where $\langle\cdot,\cdot\rangle_{2D}$ is the regular euclidean inner product in two dimensions, and $M_1, M_2 \in \Lambda_N \setminus \{0\}$, [21, Ch. 12].

The computation of the discrete derivative is performed by first dividing the EPI into blocks $l_{i,j}$ with fixed dimensions $(n_1+1)\times(n_2+1)$. Let $\beta_{n,m}(i,j)$ be the coefficient obtained from (9) with $I^{(y)}(x,s) = l_{i,j}$, and $\mathcal{B}_{n,m}$ be the matrix with the coefficients $\beta_{n.m}(i,j)$ for all pixels of $I^{(y)}(x,s)$. One can obtain $\mathcal{B}_{n,m}$ from the convolution $\mathcal{B}_{n,m} = I^{(y)}(x,s) * \mathcal{C}_{n_1,n_2}(n,m)$, where $\mathcal{C}n_1,n_2(n,m)$ is a matrix that depends only on the size of each window, and is computed as $\mathcal{C}_{n_1,n_2}(n,m) = \bar{a}_n^T\bar{b}_m$, where $\bar{a}_n$ is the row vector $[\mathbf{a}_n(0),\ldots,\mathbf{a}_n(n_1)]$ and $\bar{b}_m$ is the row vector $[\mathbf{b}_n m(0),\ldots,\mathbf{b}_m(n_2)]$, with:

$$\mathbf{a}_n(v) = \frac{k_n^{\alpha_1}(v,n_1)w_{n_1,\alpha_1}(v)}{||k_n^{\alpha_1}||} \qquad (10)$$

and

$$\mathbf{b}_m(v) = \frac{k_m^{\alpha_2}(v,n_2)w_{n_2,\alpha_2}(v)}{||k_m^{\alpha_2}||}. \qquad (11)$$

For each pixel (i,j) of the EPI $I^{(y)}(x,s)$, the discrete partial derivative for the approximation $\mathbf{P}_{M_1,M_2}$ is computed only for the region $l_{i,j}$, resulting in two matrices P_x and P_s with size equal to the EPI. From (6) and (7) it is possible to prove that $P_x = \mathcal{B}_{1,0} - \mathcal{B}_{1,2} + \frac{3}{2}(\mathcal{B}_{3,2} - \mathcal{B}_{3,0})$ and $P_s = \mathcal{B}_{0,1} - \mathcal{B}_{2,1} + \frac{3}{2}(\mathcal{B}_{2,3} - \mathcal{B}_{0,3})$.

The remaining goal is to create an edge map $e_t^{(y)}(x,s)$ from P_x and P_s. Firstly, an edge strength matrix is calculated as $\mathbf{G}(i,j) = \sqrt{P_x^2(i,j) + P_s^2(i,j)}$. Secondly, the first level of an adaptive threshold is used to find strong edge points:

$$\tau_{h1} = \text{mean}(\mathbf{G}(i,j)) + k\cdot\text{std}(\mathbf{G}(i,j)), \qquad (12)$$

with $k \in \mathbb{R}^+$. If $G(i,j) > \tau_{h1}$, (i,j) the edge is considered a strong one. The second threshold τ_{h2} is computed similarly to τ_{h1}, but using only the points (i_1,j_1) for which $mean(\mathbf{G}(i,j)) < \mathbf{G}(i_1,j_1) < \tau_{h1}$. All points for which $G(i,j) > \tau_{h2}$ are considered weak edge points. If a pixel (i,j) is a strong edge, $e_t^{(y)}(i,j) = 1$. If a pixel (i,j) is a weak edge, then $e_t^{(y)}(i,j) = 1$ if, and only if, it is adjacent to a strong edge. The resulting edge map is rough, containing thick and discontinuous lines. This is corrected by applying consecutive dilation and thin morphological operations.

B. Detection of Enlarged Regions

The texture edge maps $e_t^{(y)}(x,s)$, determined as described in the previous section, are used to improve the localization accuracy of object edges in the disparity map obtained with the structure tensor. For this purpose, the position of texture edges (points with $e_t^{(y)}(x,s) = 1$) must be compared with the position of disparity edges (points with $e_d^{(y)}(x,s) = 1$), where $e_d^{(y)}(x,s)$ is a disparity edge map.

Unlike regular images, it is not expected that neighboring points representing the same surface have a low gradient in

Algorithm 1: Detection of enlarged regions

```
Function detectEnlargedRegion(epi,disp,rel) is
    (epiHeight, epiWidth) ← size(epi);
    s_mid ← epiHeight/2;
    tEdges ← krawtchoukEdgeDetection(epi);
    dEdges ← dispEdgeDetection(disp);
    i ← 0;
    k ← 0;
    while i < epiWidth do
        if tEdges[s_mid, i] then
            for j ← i − ω to i + ω do
                if dEdges[s_mid, j] then
                    enlRegion[k] ← (i, j);
                    i ← j;
                    k ← k + 1;
                    break;
                end
            end
        else
            i ← i + 1;
        end
    end
return enlRegion
```

disparity maps. In fact, any surface at an angle near $90°$ to the camera would present a very high disparity gradient. Therefore, in such cases neither Krawtchouk polynomial based methods (as in [11]) nor other gradient-based methods (as in [10]), provide the ideal behavior.

A simple alternative is to use the Laplacian of the disparity map $\nabla^2 d^{(y)}(x,s)$, creating the binary edge map:

$$e_d^{(y)}(x,s) = \begin{cases} 0 \text{ if } \nabla^2 d^{(y)}(x,s) & \leq th \\ 1 \text{ if } \nabla^2 d^{(y)}(x,s) & > th \end{cases}, \tag{13}$$

where th is a predefined threshold.

To solve the matching problem it is necessary to first define a vicinity region N around each x_0 s.t. $e_t^{(y)}(s_0, x_0) = 1$:

$$N = \{x : |x - x_0| < w\}, \ \forall x \in \mathbb{N}, \tag{14}$$

where w defines the size of the vicinity.

For a given texture edge, potential matching candidates exist if the following condition is verified:

$$\exists\, x_0' \in N \,|\, e_d^{(y)}(s_0, x_0') = 1. \tag{15}$$

Matching candidates are only considered true matches if the following constraint is respected:

$$d^{(y)}\left(s_{mid}, x_0' + \frac{(x_0 - x_0')}{|x_0 - x_0'|}\right) < d^{(y)}(s_{mid}, x_0). \tag{16}$$

This constraint ensures that the disparity of the foreground region is larger than the disparity of the background region.

If two edges are considered true matches, the set of x coordinates E of the region between the matched edges is defined as:

$$E = \begin{cases} \{x : x \leq x_0'\}, \text{ if } x, x_0' \geq x_0 \\ \{x : x \geq x_0'\}, \text{ if } x, x_0' < x_0 \end{cases}. \tag{17}$$

Thus, $\forall e \in E$, $d^{(y)}(s_0, e)$ contains disparity values relative to the foreground region, when it should contain the values relative to the background region. The procedure to find the boundaries of this region E is illustrated in Algorithm 1.

C. Enhancement of Object Boundaries

Algorithm 2: Enhancement of object boundaries.

```
Function enhObjBound(disp,rel,boundaries) is
    for boundary in boundaries do
        (dEdge, tEdge) ← boundary;
        (dHeight, dWidth) ← size(disp);
        s_mid ← dHeight/2;
        if dEdge<tEdge then
            range ← [dEdge, ..., tEdge];
            bgRange ← [dEdge − p, ..., dEdge];
        else
            range ← [tEdge, ..., dEdge];
            bgRange ← [dEdge, ..., dEdge + p];
        end
        bgRange ← excludeUnreliable(rel, bgRange);
        bgValue ← median(disp[s_mid, bgRange]);
        disp[s_mid, range] ← bgValue;
    end
return disparity
```

Having found the pixels of the disparity map affected by silhouette enlargement, it is now the goal to correct these values to improve the accuracy of the disparity map. This is done by value replacement in the disparity map so that:

$$d^{(y)}(s_0, e) = \text{median}(d^{(y)}(s_0, b)), \forall b \in B, \tag{18}$$

where B represents the region containing pixels in the aforementioned background region.

Unlike in [9], where the region between the matching disparity edge and the next edge found in the disparity map is used, here only pixels with reliable disparity values within p pixels of the enlarged disparity edge are considered so that

$$B = \{x_0, \ldots, x_0 + (p-1)\} \cup \{x : r^{(y)}(x,s) > 0.7\}. \tag{19}$$

This solves three major issues. Firstly, it reduces the effect of false positive edges in $e_d^{(y)}(x,s)$. Secondly, when the second disparity edge is correctly found, the median of the disparities on the entire section of an inclined surface will be a bad substitute for the values at the end of the surface that's near the occlusion of the object, considering only values closer to the region to be corrected provides much more accurate results. Additionally, to aid in the robustness of the algorithm, if $B = \varnothing$ due to low reliability values, the region is left as a

non-sense value (ie. NaN). The entire procedure is illustrated in Algorithm 2.

In the disparity map $d(x, y)$ of the chosen view s_0, t_0 these values are inpainted using a Laplacian based interpolator. While achieving worse results than the replacement in (19), the result for these cases is still overwhelmingly better than the uncorrected disparity map when compared to a ground truth disparity.

D. Optimization Step

Having determined a corrected local estimation of the disparity, it is necessary to take into account constraints between points, by propagating information with a higher reliability to pixels with lower reliability. A Laplacian-based image matting approach [22] is adapted for this purpose. The optimization process is here shown for a single view of the LF, yet an extension to the entire LF is straightforward.

An energy function is defined by combining a smoothing term ($\boldsymbol{d}^T L \boldsymbol{d}$) and a data term ($\lambda(\boldsymbol{d}–\widetilde{\boldsymbol{d}})^T C(\boldsymbol{d}–\widetilde{\boldsymbol{d}})$), where $\boldsymbol{d}$ represents a vectorized version of the estimated disparity map and $\boldsymbol{d}$ is the optimal vectorized disparity map:

$$J(\boldsymbol{d}) = \boldsymbol{d}^T L \boldsymbol{d} + \lambda(\boldsymbol{d}–\widetilde{\boldsymbol{d}})^T C(\boldsymbol{d}–\widetilde{\boldsymbol{d}}), \tag{20}$$

with $\boldsymbol{d}$ and $\widetilde{\boldsymbol{d}}$ being $N \times 1$ vectors, where N represents the number of pixels in each view (i.e. $N = \text{Height} \times \text{Width}$), and L represents an affinity matrix, whose $N \times N$ elements can be understood as a measure of the affinity between each pixel and the rest of the pixels on the image. Furthermore, λ is a global weight of the data term determining the strength of the smoothing operation ($\lambda = 5$, in [22]) and C is a diagonal $N \times N$ matrix, whose diagonal entries are the reliability values $r^{(y)}(x, s)$ inputted to the algorithm.

IV. Results

The proposed method was evaluated using the commonly used HCI 4D LF dataset, made available in [23], which provides LFs with 9×9 views and a ground truth disparity map. This training dataset is used as it provides a ground truth for photo-realistic images.

The achieved results are presented in Table I, in terms of the mean squared error (MSE) between the estimated disparity maps and the ground truth, by using different edge detection algorithms. The first method (Plain) generates the maps using the structure tensor and a matting-based optimization. The second method (Canny) estimates the maps using the silhouette enhancement algorithm presented in [9]. The proposed method (Prop.) estimates the disparity maps as described in Section III. All methods use the same matting-based optimization, which allows to highlight only the improvements related to the enhancement of object boundaries. The results show a relative improvement of up to 34.15%, in the case of the *Cotton* image, while also obtaining meaningful reductions in MSE for the Boxes and Cotton images.

While the impact on the mean squared error is considerable, another significant contribution of the proposed method is the reduction of artefacts around image boundaries. These artefacts occur whenever there are incomplete edge segments (false negatives) in the estimated object edge map. Figure 2 depict some details from the *Cotton* image using the *Canny* edge detector and the proposed edge detection method. In the disparity maps on the left, referring to those obtained using the Canny method, it is possible to observe annoying artefacts, where a large number of false negatives resulted in no correction being applied. These artefacts are less visible in right-most disparity maps, which were obtained using the proposed method for silhouette improvement.

TABLE I
MSE ($\times 100$) FOR DIFFERENT EDGE DETECTION ALGORITHMS.

Images	Plain	Canny	Prop.
Boxes	9.56	9.31	**8.33**
Cotton	2.21	0.79	**0.52**
Dino	1.19	0.63	**0.55**
Sideboard	1.91	1.21	**1.21**

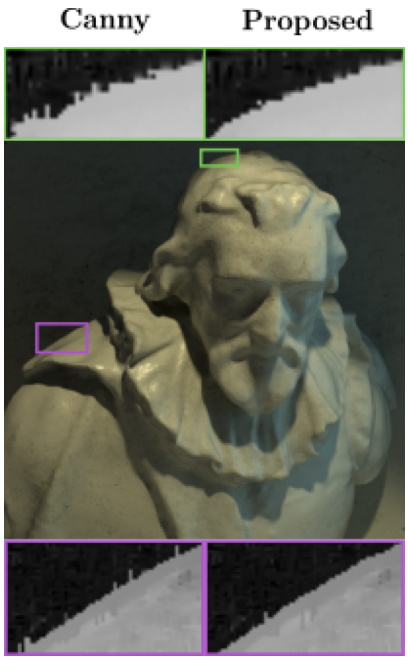

Fig. 2. Comparison of border artefacts on the disparity map.

The proposed approach has also been evaluated against other state-of-the-art methods for disparity estimation, as shown in Table II. The Spinning Parallelogram Operator (SPO) [24] is among the best non-supervised algorithms in most metrics across the HCI dataset. The EPI2 method [7] is a structure tensor-based method that uses a more robust optimisation step, instead of an explicit algorithm for better silhouette localization. For all images, the proposed method achieves a better MSE than the EPI2 method. It also shows to perform better than the SPO algorithm, on average, particularly in half of the tested images, while significantly outperforming SPO for the *Cotton* image. The best results are achieved when the method is applied to images with fewer but larger objects

with well defined edges.

TABLE II
MSE ($\times 100$) COMPARISON WITH THE STATE-OF-THE-ART METHODS.

Images	Plain	SPO	EPI2	Prop.
Boxes	9.56	9.11	10.93	**8.33**
Cotton	2.21	1.31	4.32	**0.52**
Dino	1.19	**0.31**	2.08	0.55
Sideboard	1.91	**1.02**	4.65	1.21
Average	3.72	2.94	5.495	**2.65**

V. CONCLUSIONS

In this work, the locality of silhouettes in LF disparity maps, obtained from the structure tensor, is improved with the use of the proposed edge detecting method based on Krawtchouk polynomials. The edge detecting algorithm is positively compared with the Canny edge detector, showing noticeable improvements in terms of mean square error and a large reduction in the number of visual impact artefacts. When compared to other state-of-the-art algorithms using the HCI image dataset, the proposed algorithm proves to perform better than other methods based on the structure tensor, namely Plain and EPI2, and also outperforming the Spinning Parallelogram Operator method.

REFERENCES

[1] S. McDonagh, R. B. Fisher, and J. Rees, "Using 3d information for classification of non-melanoma skin lesions," in *Medical Image Understanding and Analysis*, no. 164-168. Dundee, UK: BMVA Press, Janary 2008.

[2] P. Schelkens, P. Astola, E. A. B. da Silva, C. Pagliari, C. Perra, I. Tabus, and O. Watanabe, "JPEG Pleno light field coding technologies," in *Applications of Digital Image Processing*, A. G. Tescher and T. Ebrahimi, Eds., vol. 11137, International Society for Optics and Photonics. San Diego, California, United States: SPIE, September 2019, pp. 391 – 401. [Online]. Available: https://doi.org/10.1117/12.2532049

[3] M. W. Tao, S. Hadap, J. Malik, and R. Ramamoorthi, "Depth from combining defocus and correspondence using light-field cameras," *Proceedings of the IEEE International Conference on Computer Vision*, pp. 673–680, 2013.

[4] H. Lin, C. Chen, S. Bing Kang, and J. Yu, "Depth recovery from light field using focal stack symmetry," in *Proceedings of the International Conference on Computer Vision*, Santiago, Chile, 2015, pp. 3451–3458.

[5] W. Williem, I. K. Park, and K. M. Lee, "Robust light field depth estimation using occlusion-noise aware data costs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–1, August 2018.

[6] H. G. Jeon, J. Park, G. Choe, J. Park, Y. Bok, Y. W. Tai, and I. S. Kweon, "Accurate depth map estimation from a lenslet light field camera," in *IEEE Conference on Computer Vision and Pattern Recognition*, Boston, USA, June 2015, pp. 1547–1555.

[7] S. Wanner and B. Goldluecke, "Globally consistent depth labeling of 4d light fields," in *IEEE Conference on Computer Vision and Pattern Recognition*, Providence, RI, USA, June 2012, pp. 41–48.

[8] J. Li, M. Lu, and Z. Li, "Continuous depth map reconstruction from light fields," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 3257–3265, November 2015.

[9] R. Lourenco, P. A. A. Assuncao, L. M. N. Tavora, R. Fonseca-Pinto, and S. M. M. Faria, "Silhouette enhancement in light field disparity estimation using the structure tensor," in *25th IEEE International Conference on Image Processing*, Athens, Greece, October 2018, pp. 2580–2584.

[10] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, Nov 1986.

[11] D. Rivero-Castillo, H. Pijeira, and P. A. A. Assuncao, "Edge detection based on krawtchouk polynomials," *Journal of Computational and Applied Mathematics*, vol. 284, pp. 244–250, August 2015.

[12] M. Levoy and P. Hanrahan, "Light field rendering," in *Proceedings of the Conference on Computer Graphics and Interactive Techniques*. New York, NY, USA: ACM, 1996, pp. 31–42. [Online]. Available: http://doi.acm.org/10.1145/237170.237199

[13] R. C. Bolles, H. H. Baker, and D. H. Marimont, "Epipolar-plane image analysis: An approach to determining structure from motion," *International Journal of Computer Vision*, p. 7–55, March 1987.

[14] S. Wanner, "Orientation Analysis in 4D Light Fields," Ph.D. dissertation, Universitat Heidelberg, Germany, 2014.

[15] J. Bigun, "Optimal orientation detection of linear symmetry," in *IEEE First International Conf. on Computer Vision*, London, Great Britain, June 1987, pp. 433–438.

[16] T. S. Chihara, *An Introduction to Orthogonal Polynomials*. New York, NY, USA: Gordon and Breach, 1978.

[17] G. Szegő, *Orthogonal Polynomials*. Providence, RI, USA: Amer. Math. Soc. Colloq. Publ., 1975, vol. 23.

[18] A. Nikiforov, S. Suslov, and V. Uvarov, *Classical Orthogonal Polynomials of a Discrete Variable*. New York, NY, USA: Springer-Verlag, 1991.

[19] J. Baik, T. Kriecherbauer, K. McLaughlin, and P. Miller, *Orthogonal Polynomials*. Princeton, NJ, USA: Princeton Univ. Press, 2007.

[20] C. Dunkl and Y. Xu, *Orthogonal Polynomials of several Variables*. Cambridge: Cambridge Univ Press, 2001.

[21] J. Rice, *The approximations of functions, vol 2: Nonlinear and multivariate theory*. London: Addison-Wesley Publ. Co., Reading Mass, 1969.

[22] A. Levin, D. Lischinski, and Y. Weiss, "A Closed Form Solution to Natural Image Matting," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 2, pp. 228–242, December 2008.

[23] O. Johannsen, K. Honauer, B. Goldluecke, A. Alperovich, F. Battisti, Y. Bok, M. Brizzi, M. Carli, G. Choe, M. Diebold, M. Gutsche, H. Jeon, I. S. Kweon, J. Park, J. Park, H. Schilling, H. Sheng, L. Si, M. Strecke, A. Sulc, Y. Tai, Q. Wang, T. Wang, S. Wanner, Z. Xiong, J. Yu, S. Zhang, and H. Zhu, "A taxonomy and evaluation of dense light field depth estimation algorithms," in *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Honolulu, USA, July 2017, pp. 1795–1812.

[24] S. Zhang, H. Sheng, C. Li, J. Zhang, and Z. Xiong, "Robust depth estimation for light field via spinning parallelogram operator," *Computer Vision and Image Understanding*, vol. 145, pp. 148–159, April 2016.