Fuzzy Multi-Criteria Decision Making in Stereovision Matching for Fish-Eye Lenses in Forest Analysis

P.J. Herrera[1], G. Pajares[1], M. Guijarro[2], J.J. Ruz[3], and J.M. De la Cruz[3]

[1] Dpto. Ingeniería del Software e Inteligencia Artificial, Facultad de Informática,
Universidad Complutense, 28040 Madrid, Spain
pjherrera@pdi.ucm.es, pajares@fdi.ucm.es
[2] Centro Superior de Estudios Felipe II, Ingeniería Técnica en informática de Sistemas
28300 Aranjuez, Madrid, Spain
mguijarro@cesfelipesegundo.com
[3] Dpto. Arquitectura Computadores y Automática, Facultad de Informática, Universidad
Complutense, 28040 Madrid, Spain
{jjruz,jmcruz}@dacya.ucm.es

Abstract. This paper describes a novel stereovision matching approach based on omni-directional images obtained with fish-eye lenses in forest environments. The goal is to obtain a disparity map as a previous step for determining the volume of wood in the imaged area. The interest is focused on the trunks of the trees. Due to the irregular distribution of the trunks, the most suitable features are the pixels. A set of six attributes is used for establishing the matching between the pixels in both images of each stereo pair analysed. The final decision about the matched pixels is taken based on a well tested Fuzzy Multi-Criteria Decision Making approach, where the attributes determine the criteria and the potential matches in one image of the stereo pair for a given pixel in the other one determine the alternatives. The application of this decision making approach makes the main finding of the paper. The full procedure is based on the application of three well known matching constraints. The proposed approach is compared favourably against the usage of simple features.

Keywords: Fish-eye stereo vision, Stereovision matching, omni-directional forest images, fuzzy Multi-Criteria Decision Making.

1 Introduction

One important task in forests maintenance is to determine the volume of wood in an area for different purposes, including the control of growth of the trees. This task can be carried out by stereovision systems. Fish-eye lenses allow imaging a large sector of the surrounding space with omni-directional vision. This justifies its use.

According to [1] we can view the classical problem of stereo analysis as consisting of the following steps: image acquisition, camera modelling, feature acquisition, image matching, depth determination and interpolation. The key step is that of image matching. This is the process of identifying the corresponding points in two images that are cast by the same physical point in the 3-D space. This paper is devoted solely
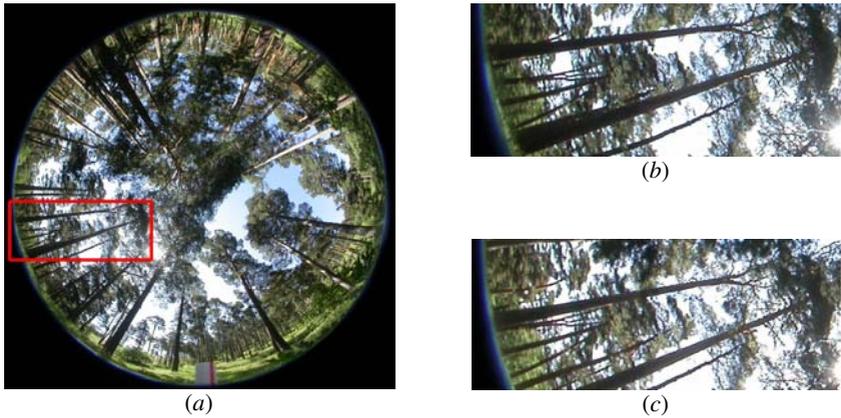
Fig. 1. (*a*) Omni-directional left image; (*b*) left expanded area; (*c*) corresponding right expanded area

to the matching one. Two sorts of techniques have been used for matching: area-based and feature based [2].

Area-based stereo techniques [3] use correlation between brightness (intensities) patterns in the local neighbourhood of a pixel in one image with brightness patterns in the local neighbourhood of the other image. Also statistical textures can be considered under this category. Feature-based methods [4] use set of pixels with similar attributes, colour, gradient (module and direction) or Laplacian. These are the six attributes available to be used in the proposed matching procedure.

Figure 1(*a*) displays one omni-directional image (let's say the left one) of the stereo pair captured with a fisheye lens. Figure 1(*b*) displays the signed and expanded area on Figure 1(*a*). In Figure 1(*c*) the corresponding area in the right image of the stereo pair is displayed. Due to the different locations of the tree's crowns there exists an important lighting variability between both areas; this makes the matching process a difficult task. This is applicable for the whole image.

The following three stereovision constraints can be applied for solving the matching problem. *Epipolar*: derived from the system geometry, given a pixel in one image its correspondence will be on the epipolar line. *Similarity*: matched pixels display similar attributes. *Uniqueness*: a pixel in the left image must be matched to a unique pixel in the right one.

Given a pixel in the left image, we apply the epipolar constraint for determining a list of candidates, which are potential matches, in the right image. Each candidate becomes an alternative for the first pixel. We also apply the similarity constraint based on the six attributes, obtaining a set of criteria. Through the fuzzy Multi-Criteria Decision Making (MCDM) approach, a unique candidate is selected as the correct match among the list of candidates. This implies the application of the uniqueness constraint. The matching through the MCDM makes the main contribution of the paper. The proposed approach is compared favourably against the usage of individual area-based and feature-based correspondence techniques.

This work is organized as follows. Section 2 describes the design of the matching process; including a brief overview of the MCDM paradigm. Section 3 describes the

results obtained by using the combined MCDM approach, and comparing these results with those obtained by applying each individual strategy. Section 4 presents the conclusions and future work.

2   Design of the Matching Process

2.1   Epipolar: System Geometry

Figure 2 displays the stereo vision system geometry [5]. The 3D object point P with world coordinates with respect to the systems (X_1, Y_1, Z_1) and (X_2, Y_2, Z_2) is imaged as (x_{i1}, y_{i1}) and (x_{i2}, y_{i2}) in image-1 and image-2 respectively in coordinates of the image system; α_1 and α_2 are the angles of incidence of the rays from P; y_{12} is the baseline measuring the distance between the optical axes in both cameras along the y-axes; r is the distance between image point and optical axis; R is the image radius, identical in both images.

According to [6], the following geometrical relations can be established,

$$r = \sqrt{x_{i1}^2 + y_{i1}^2} \; ; \; \alpha_1 = (r90^\circ)/R \; ; \; \beta = tg^{-1}\left(y_{i1}/x_{i1}\right) \tag{1}$$

Now the problem is that the 3D world coordinates (X_1, Y_1, Z_1) are unknown. They can be estimated by varying the distance d as follows,

$$X_1 = d\cos\beta; \quad Y_1 = d\sin\beta; \quad Z_1 = \sqrt{X_1^2 + Y_1^2}\Big/\tan\alpha_1 \tag{2}$$
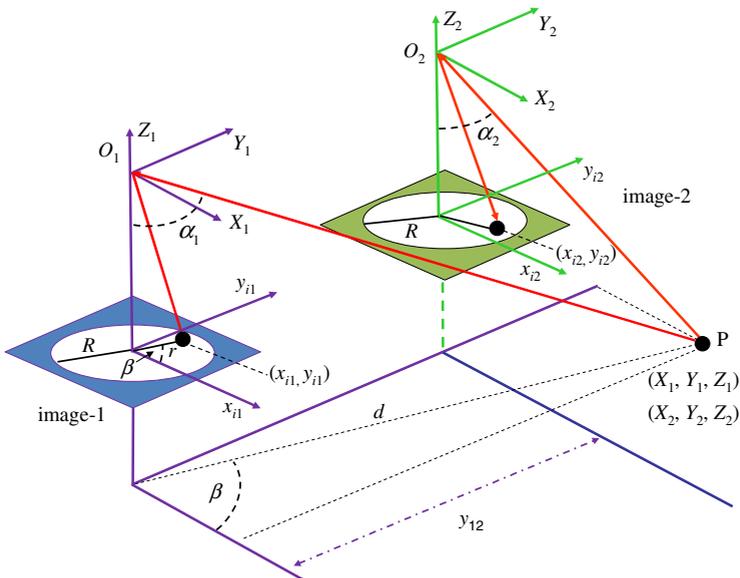

Fig. 2. Geometric projections and relations for the fish-eye based stereo vision system

From (2) we transform the world coordinates in the system $O_1X_1Y_1Z_1$ to the world coordinates in the system $O_2X_2Y_2Z_2$ taking into account the baseline as follows:

$$X_2 = X_1; \quad Y_2 = Y_1 + y_{12}; \quad Z_2 = Z_1 \tag{3}$$

Assuming no lenses radial distortion, we can find the imaged coordinates of the 3D point in image-2 as [6],

$$x_{i2} = \frac{2R\arctan\left(\sqrt{X_2^2 + Y_2^2}\Big/Z_2\right)}{\pi\sqrt{\left(Y_2/X_2\right)^2 + 1}}; \quad y_{i2} = \frac{2R\arctan\left(\sqrt{X_2^2 + Y_2^2}\Big/Z_2\right)}{\pi\sqrt{\left(X_2/Y_2\right)^2 + 1}} \tag{4}$$

Using only a camera, we capture a unique image and the 3D points belonging to the line $\overline{O_1P}$, are all imaged in the unique point (x_{i1}, y_{i1}). So, the 3D coordinates with a unique camera cannot be obtained. When we try to match the imaged point (x_{i1}, y_{i1}) into the image-2 we follow the epipolar line, i.e. the projection of $\overline{O_1P}$ over the image-2. This is equivalent to vary the parameter d in the 3-D space. So, given the imaged point (x_{i1}, y_{i1}) in the image-1 (left) and following the epipolar line, we obtain a list of m potential corresponding candidates represented by (x_{i2}, y_{i2}) in the image-2 (right).

2.2 Similarity: Attributes for Area and Feature-Based

Each pixel l in the left image is characterized by its attributes; one of such attributes is denoted as A_l. In the same way, each candidate i in the list of m candidates is described by identical attributes, A_i. So, we can compute differences between attributes of the same type A, obtaining a similarity measure for each attribute as,

$$s_{iA} = \left(1 + |A_l - A_i|\right)^{-1}; \quad i = 1, ..., m \tag{5}$$

$s_{iA} \in [0,1]$, $s_{iA} = 0$ if the difference between attributes is large enough (minimum similarity), otherwise if they are equal ($s_{iA} = 1$, maximum similarity).

As mentioned before, in this paper we use the following six attributes for describing each pixel (feature): a) correlation; b) texture; c) colour; d) gradient magnitude; e) gradient direction and f) Laplacian. Both first ones are area-based computed on a 3×3 neighbourhood around each pixel through the correlation coefficient [7] and standard deviation [8]. The four remaining ones are considered as feature-based [4]. The colour involves the three red-green-blue spectral components (R,G,B) and the absolute value in the equation (5) is extended as: $|A_l - A_i| = \sum_H |H_l - H_i|$, $H =$ R,G,B.

Gradient (magnitude and direction) and Laplacian are computed by applying the first and second derivatives [8], over the intensity image after its transformation from the RGB plane to the HSI (hue, saturation, intensity) one. Given a pixel in the left image and the set of m candidates in the right one, we compute the following similarity measures for each attribute A: s_{ia} (correlation), s_{ib} (colour), s_{ic} (texture), s_{id} (gradient

magnitude), s_{ie} (gradient direction) and s_{if} (Laplacian). The identifiers in the sub indices identify the attributes according to the above assignments.

2.3 Uniqueness: Applying the Fuzzy MCDM Paradigm

Now we must match each pixel l in the left image with the best of the potential candidates (uniqueness). This is a decision based on the fuzzy MCDM paradigm. Before, we need to define the following concepts [9]: a) a triangular fuzzy number u as a triplet (u_1, u_2, u_3) and b) a distance between two triangular fuzzy numbers u and z as follows,

$$d(u,z) = \left\{ \left[(u_1 - z_1)^2 + (u_2 - z_2)^2 + (u_3 - z_3)^2 \right] \Big/ 3 \right\}^{1/2} \tag{6}$$

Taking into account the set of the six similarity measures, equation (5), we can group them into two groups C_1 and C_2, each with three values, namely: $C_1 = \{c_{11}, c_{12}, c_{13}\}$ and $C_2 = \{c_{21}, c_{22}, c_{23}\}$. Each group defines a similarity criterion, i.e. we have available two criteria for making the decision C_j $(j = 1,…,n)$, where n is 2 in this approach. Under the fuzzy MCDM paradigm [9, 10] and assuming the m candidates, this can be formulated as the choice of the best alternative M_i $(i = 1,…, m)$; where now each alternative represents a candidate. The MCDM problem can be expressed in matrix format as follows,

$$D = \left[x_{ij} \right]_{m \times n} ; \quad W = \left[w_j \right]_{1 \times n} ; \quad i = 1,...,m; \; j = 1,...,n. \tag{7}$$

D is the decision matrix where x_{ij} is the rating of alternative M_i with respect to the criterion C_j; w_j is the weight assigned to the criterion C_j. We consider the similarity measures as triangular fuzzy numbers so that $x_{i1} = (s_{ia}, s_{ib}, s_{ic})$ and $x_{i2} = (s_{id}, s_{ie}, s_{if})$. The weights associated with every similarity measure are respectively $w_1 = (w_a, w_b, w_c)$ and $w_2 = (w_d, w_e, w_f)$, estimated according to its relative importance, as described later in the section 3a. Without loss of generality, the values in x_{i1}, x_{i2} are ordered so that $s_{ia} \le s_{ib} \le s_{ic}$ and also for $s_{id} \le s_{ie} \le s_{if}$. This arrangement is applied to the list of the m candidates. Therefore, we can obtain the normalized fuzzy decision matrix as follows,

$$R = \left[r_{ij} \right]_{m \times n} ; \quad r_{i1} = \left(\frac{s_{ia}}{s_{ic}^*}, \frac{s_{ib}}{s_{ic}^*}, \frac{s_{ic}}{s_{ic}^*} \right) ; \quad r_{i2} = \left(\frac{s_{id}}{s_{if}^*}, \frac{s_{ie}}{s_{if}^*}, \frac{s_{if}}{s_{if}^*} \right) \tag{8}$$

where $s_{ic}^* = \max_i \{ s_{ic} \}$ and $s_{if}^* = \max_i \{ s_{if} \}$. This preserves the property that the ranges of normalized triangular fuzzy numbers belong to the interval [0, +1]. Considering the importance assigned to every criterion, we can construct the weighted normalized fuzzy decision matrix in equation (9).

$$V = \left[v_{ij} \right]_{m \times n} \quad \text{where} \quad v_{i1} = r_{i1} w_{j1}; \quad v_{i2} = r_{i2} w_{j2} \tag{9}$$

From V, we know that the elements $v_{ij}, \forall ij$ are normalized positive triangular fuzzy numbers ranging in the closed interval $[0, +1]$. Then we can define the fuzzy positive-ideal solution $p^+ = (1,1,1)$ and the fuzzy negative-ideal solution $p^- = (0,0,0)$. The distances for each alternative can be calculated as follows,

$$d_i^+ = d(v_{i1}, p^+) + d(v_{i2}, p^+) \quad \text{and} \quad d_i^- = d(v_{i1}, p^-) + d(v_{i2}, p^-) \tag{10}$$

where $d(\cdot,\cdot)$ is the distance measured between two fuzzy numbers, defined in the equation (6). According to [9] a closeness coefficient is defined to determine the ranking order of all alternatives, once both d_i^+ and d_i^- for each alternative have been computed. This coefficient is,

$$CC_i = d_i^- / (d_i^+ + d_i^-) \tag{11}$$

Obviously, an alternative M_i is closer to the fuzzy ideal solution and farther from the fuzzy negative solution as CC_i approaches to $+1$. So, given l its best match i is that with CC_i maximum.

3 Results

The system is based on the scheme in figure 2, with a baseline of 1 meter. The cameras are equipped each one with Nikon FC-E8 fisheye lens, with an angle of 183°. The valid colour images in the circle contain 6586205 pixels.

   The tests have been carried out with twenty pairs of stereo images. We have four of them for computing the relevance of each criterion, from which the fuzzy weights are obtained. At a second stage, we apply the MCDM approach pixel by pixel for the remainder sixteen stereo pairs.

   Our interest consists of determining the disparity of the trees trunks located in an area of 25 m^2 around the stereo vision system. The disparity is the absolute difference value in sexagesimal degrees, taking into account the imaged circle, between the pixel in the left image and its matched pixel in the right one. We have available the information of disparities provided by the end users. Thus, for each pixel in a trunk we know its correct disparity value according to this expert knowledge; which allows us to compute the percentage of error. For each one of the sixteen stereo images used for testing, we compute the disparity error for the trunks and then average these errors among the sixteen pairs of stereo images.

a) Computing the relevance for each criterion

Given both available stereo images for this purpose, for each pixel in the left images, we compute the disparity with respect its matched pixel in the right ones, but considering each one of the six attributes separately through the equation (5). So, we compute the averaged percentage of error for both stereo images and for each attribute, based on the expert knowledge available about the disparities in the trunks. They are: $p_a = 28$ (correlation), $p_b = 10$ (colour), $p_c = 14$ (texture), $p_d = 9$ (gradient magnitude),

$p_e = 30$ (gradient direction) and $p_f = 27$ (Laplacian). So, each weight is computed as $w_h = p_h / \sum_k p_k$, h, k = a,b,c,d,e,f. Finally, $w_1 \approx 10^{-3}(150,187,179)$ and $w_2 \approx 10^{-3}(189,145,152)$. As one can see, the most relevant attribute is the gradient magnitude.

b) MCDM performance

As before, for each pixel in each one of the sixteen stereo images, available for testing, we obtain its disparity considering the six attributes separtley by applying the equation (5) and a maximum similarity criterion among the *m* candidates and also by applying the MCDM approach based on the closeness coefficient, equation (11).

Figures 3(*a*) and 3(*b*) are the same that Figures 1(*a*) and 1(*b*) respectively. Figure 3(*c*) displays the disparity map obtained by the MCDM approach in the area. The colour bar shows the disparity level values according to the colour. Table 1 displays the averaged percentage of error based on the similarity for the six individual attributes used separately, identified as: (s_a, s_b, s_c, s_d, s_e, s_f). The averaged percentage of error obtained with the MCDM approach is also displayed. The standard deviations are also included.

From results in table 1, one can see that MCDM outperforms the individual similarity based approaches. This means that the combination of similarities between attributes improve the results obtained by using similarities separately. The best individual similarity results are obtained through the similarities provided by the gradient magnitude attribute (s_d). This implies that it is the most relevant attribute. This agrees to its relevance obtained previosly, as it has turned out to be the most relevant attribute.
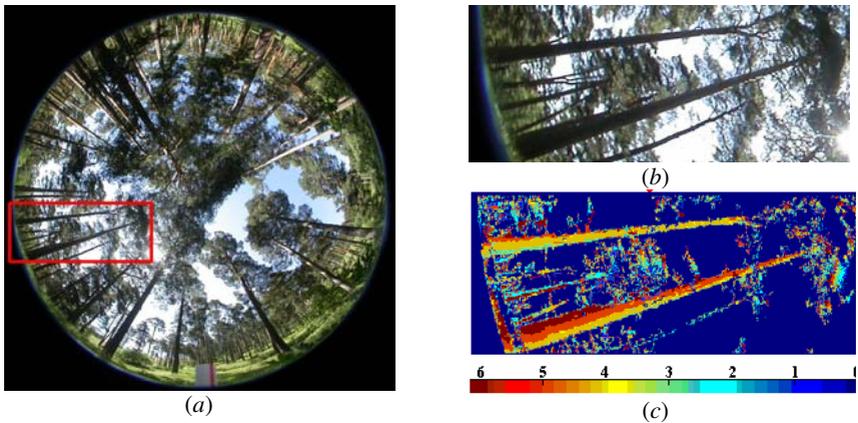
Fig. 3. (*a*) Left image; (*b*) expanded area; (*c*) disparity map obtained by the MCDM approach

Table 1. Averaged percentage of errors and standard deviations obtained through maximum similarity criteria for each attribute separately and the MCDM decision making approach

Averaged percentage of error and standard deviations													
s_a		s_b		s_c		s_d		s_e		s_f		MCDM	
%	σ	%	σ	%	σ	%	σ	%	σ	%	σ	%	σ
30	2.9	16	1.3	18	1.7	14	1.1	35	3.6	32	3.1	9	0.9

4   Concluding Remarks

In this paper we have proposed a method for stereovision matching, with omni-directional images, in a system equipped with fish-eye lenses. The method applies three well-known constraints (*epipolar*, *similarity* and *uniqueness*) by combining area-based and feature-based matching strategies. For each pixel in the left image, a list of possible candidates in the right image is obtained for determining its correspondence. The similarity between attributes establishes measures for the matching between the pixel and its candidates. Under the MCDM paradigm, we combine the similarities between six attributes and make a decision for choosing the unique candidate if any for a given pixel. The proposed combined strategy outperforms the methods that use similarities separately and it is compared favorably.

Although the results achieved can be considered satisfactory, they could be improved by applying additional constraints such as *smoothness* or *ordering*, which have been used for matching in conventional stereovision systems.

Acknowledgements

References

1. Barnard, S., Fishler, M.: Computational stereo. ACM Computing Surveys 14, 553–572 (1982)
2. Cochran, S.D., Medioni, G.: 3-D surface description from binocular stereo. IEEE Trans. Pattern Analysis and Machine Intelligence 14(10), 981–994 (1992)
3. Tang, L., Wu, C., Chen, Z.: Image dense matching based on region growth with adaptive window. Pattern Recognit. Letters 23, 1169–1178 (2002)
4. Lew, M.S., Huang, T.S., Wong, K.: Learning and feature selection in stereo matching. IEEE Trans. Pattern Anal. Machine Intell. 16, 869–881 (1994)
5. Abraham, S., Förstner, W.: Fish-eye-stereo calibration and epipolar rectification. Photogrammetry and Remote Sensing 59, 278–288 (2005)
6. Schwalbe, E.: Geometric modelling and calibration of fisheye lens camera systems. In: Proc. 2nd Panoramic Photogrammetry Workshop, Int. Archives of Photogrammetry and Remote Sensing, Part 5/W8, vol. 36 (2005)
7. Barnea, D.I., Silverman, H.F.: A class of algorithms for fast digital image registration. IEEE Trans. Computers 21, 179–186 (1972)
8. Pajares, G., de la Cruz, J.M.: Visión por Computador: Imágenes digitales y aplicaciones. RA-MA (2008)
9. Chen, C.T.: Extensions of the TOPSIS for group decision-making under fuzzy environment. Fuzzy Sets and Systems 114, 1–9 (2000)
10. Wang, W., Fenton, N.: Risk and confidence analysis for fuzzy multi criteria decision making. Knowledge Based Systems 19, 430–437 (2006)