Quasi-Dense Matching between Perspective and Omnidirectional Images

Lingling Lu, Yihong Wu

Quasi-Dense Matching between Perspective and Omnidirectional Images

Lingling Lu and Yihong Wu

National Laboratory of Pattern Recognition, Institute of Automation
Chinese Academy of Sciences, P.O. Box 2728, Beijing 100190, P.R. China
{lllu, yhwu}@nlpr.ia.ac.cn

Abstract. In this paper, we propose a match propagation algorithm between perspective and omnidirectional images of same scene consisting of planes, which does not require a rectification for the omnidirectional image. First, a linear transformation is introduced to identify the area containing the corresponding point candidates. Then, a geometric invariant is computed as a constraint for quasi-dense matching. Finally, combining the computed geometric invariant with a best-first strategy of Lhuillier and Quan (2002), the quasi-dense point correspondences are calculated. The experiments with real data show that the algorithm of this paper has good performance.

1 Introduction

Robot vision is an important part for a robot system. Efficient vision sensors expected as human eyes are needed when a robot faces a complex environment. Some researchers have designed vision sensors mimicking the human eye structure [1–5]. Each eye of these vision sensors consists of two cameras, of which one is used to monitor surroundings with wide-angle field of view and the other one with narrow-angle field of view is used to gaze at interest objects. The camera with wide-angle field of view could be realized by an omnidirectional camera such as a catadioptric camera or a fisheye camera. The one with narrow-angle view field could be realized by a perspective camera. This eye structure with one omnidirectional camera and one perspective camera is called perspective-omnidirectional hybrid camera system (POHC system) here.

Image matching between perspective and omnidirectional cameras is helpful for a robot to track, recognize or reconstruct interest objects [1]. Many fairly mature image matching algorithms between perspective cameras have been proposed [6–10]. There are also some image matching algorithms between omnidirectional cameras [11, 12]. However, no automatic image matching has been proposed between perspective and omnidirectional cameras.

The image matching algorithms between perspective cameras [6–10, 13–15] can't be used for POHC system directly because straightforward perspective image matching algorithms for POHC produces only very few point correspondences. In some applications like 3D reconstruction, a large quantity of high

accuracy point correspondences is needed. Sparse point correspondences can be used for obtaining quasi-dense point correspondences by match propagation or quasi-dense matching [13, 14]. Quasi-dense matching has some advantages in some applications of POHC system. For uncalibrated image sequences, it provides automatic geometry estimation as the sparse matching. Still the resulting geometric estimation is more robust and accurate. In addition, quasi-dense matching could give more points displaying objects in more detail. Simultaneously, it is more efficient than dense matching but sufficient for accurate 3D description of the scene.

Therefore, in this paper, we take a study on the match propagation between perspective and omnidirectional views. A geometric invariant between perspective and omnidirectional cameras is used as a geometric constraint. Inspired by Kannala and Brandt [15], we use a general affine model for the local transformation between the patches. During the propagation, the affine transformation matrices of new matches are achieved by using the second order intensity moments locally together with the main directions of corresponding points.

The organization of the paper is as follows. First, the existing match propagation algorithms between two traditional perspective views are described briefly in Section 2. Section 3 introduces the geometric invariant between perspective and omnidirectional cameras and gives the match propagation algorithm between perspective and omnidirectional views. The experiments are reported in Section 4, and Section 5 makes some conclusions.

2 Preliminaries

Match propagation is a technique for computing quasi-dense point correspondences between two views by propagating initial sparse point correspondences to neighboring pixels. Lhuillier and Quan [13] presented a match propagation algorithm which uses a best-first strategy and is very effective for narrow baseline views. Then, Kannala and Brandt [15] extended their method to wide baseline views. Next we describe the quasi-dense wide baseline match algorithm briefly.

The initial matches are achieved as follows: the detector is the Hessian-Affine detector [10], and the detector regions are matched using the **SIFT** descriptor [6]. Every interest point has an affine covariant region (elliptical region), therefore, the affine transformation matrices of the initial matches can be computed by using the equations of the ellipses and the main directions of the interest points.

The initial matches are called seed points. The propagation proceeds by iterating the following three steps until the list of seeds becomes empty:

(i) The seed with the best **ZNCC** (Zero-mean Normalized Cross-Correlation) score [13] is removed from the list of seed points.

(ii) New candidate matches are searched from the spatial neighborhood of the current seed.

(iii) The candidate matches which meet the geometric constraint and exceed a **ZNCC** threshold are stored in the result of the quasi-dense match and added to the list of seed points.

More precisely, in the above step (i) the **ZNCC** scores are computed not from square image patches any longer but from geometrically normalized image patches. The normalization is done as follows: a square patch of size $(2W+1) \times (2W+1)$ centered on the seed point is extract from the image which locally has a lower resolution (the local magnification factor is given by $abs(det\mathbf{A})$. $\mathbf{A}$ is the affine transformation matrix of the seed point); the corresponding area in the other image is obtained as the affinely mapped result from the square patch of the first image; the patches from both images are transformed to $(4W+1)\times(4W+1)$ windows by interpolation and then decimated to the size $(2W+1) \times (2W+1)$.

In Step (ii) of the algorithm, the neighborhood from which new matches are searched also needs to be normalized.

In Step (iii) the geometric constraint is the epipolar constraint using fundamental matrix. When new matches are added to the list of seeds they can inherit the affine transformation matrix from the current seed. The better way is to allow the adaptation of the local affine transformation during the propagation after the epipolar geometry is known. The adaptation is extracted from the second order intensity moments which determine the local affine transformation up to a rotation. The remaining rotational degree of freedom is determined from the epipolar geometry. If the second order moment matrices are $\mathbf{S}$, $\mathbf{S}'$ and the rotational matrix is $\mathbf{R}$, then the affine transformation matrix $\mathbf{A}$ is computed as follows:$\mathbf{A} = \mathbf{S}'^{0.5}\mathbf{R}\mathbf{S}^{-0.5}$.The details and the computation of S, S' can be found in [15].

Some used projective geometric concepts are as follows. A pencil of lines is a set of lines through a fixed point. The fixed point is called the vertex of the pencil. For a pencil of four coplanar lines, a projective geometric invariant (a cross ratio) can be computed from 1D line-coordinates of these lines. A pencil of planes is a set of planes in space through a fixed line, and the fixed line is called the axis of the pencil. There is a cross ratio from a pencil of four planes. This cross ratio is equal to the cross ratio of four intersecting lines of a space plane with the planes of the pencil [16].

3 Match Propagation for Perspective and Omnidirectional Views

Match propagation for perspective views use the epipolar constraint as the geometric constraint [13, 14]. We introduce a new geometric constraint for perspective and omnidirectional views. The given geometric constraint is described with the typical bracket algebra that is free of coordinate system [17]. The derivation is under the central catadioptric model, which is still in fact valid for a more general model: the locus of optical centers lies on a segment of the optical axis rather than a single viewpoint and with only radial distortion, such as fisheye cameras [18].

First, we describe the geometric constraint between perspective and omnidirectional images in detail. Then, we estimate this constraint and study a local

affine transformation for propagation. Finally, the outline of the proposed algorithm is given.

3.1 Geometric Constraint between Omnidirectional and Perspective Image Points

The omnidirectional camera model is an equivalent spherical projection defined by Geyer and Daniilidis [19] as shown in Fig. 1. We briefly recall this model here. A space point **M** is projected to a point **X** on the viewing sphere through the sphere center **O**, and then projected to **m** on the image plane through the camera viewpoint $\mathbf{O}_c$.

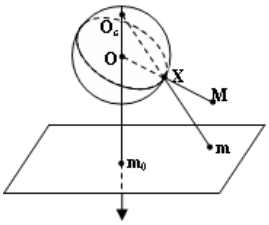

Fig. 1. The imaging process under the omnidirectional camera model

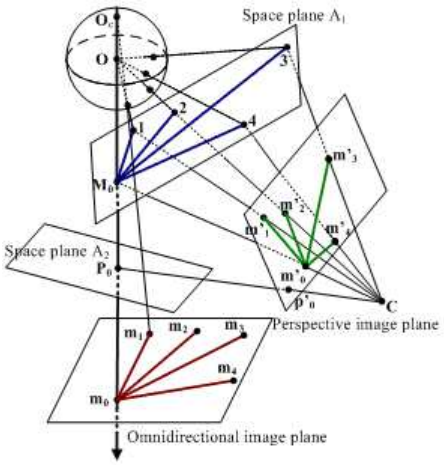

Fig. 2. Relationship between omnidirectional and perspective image points

Wu and Hu [18] presented geometric invariants between space points and image points under central catadioptric camera model. Let **1**, **2**, **3**, **4** be four coplanar space points as shown in Fig. 2 and the base plane of them is denoted as A_1. The intersection point of A_1 with the optical axis $\mathbf{OO}_c$ is denoted as $\mathbf{M}_0$. The image points of $\mathbf{M}_0$, **1**, **2**, **3**, **4** are respectively $\mathbf{m}_0$, $\mathbf{m}_1$, $\mathbf{m}_2$, $\mathbf{m}_3$, $\mathbf{m}_4$. Then the following invariant equation holds:

$$\frac{[\mathbf{13M}_0][\mathbf{24M}_0]}{[\mathbf{23M}_0][\mathbf{14M}_0]} = \frac{[\mathbf{m}_1\mathbf{m}_3\mathbf{m}_0][\mathbf{m}_2\mathbf{m}_4\mathbf{m}_0]}{[\mathbf{m}_2\mathbf{m}_3\mathbf{m}_0][\mathbf{m}_1\mathbf{m}_4\mathbf{m}_0]}. \tag{1}$$

where [ ] denotes the determinant of vectors in it, and $\mathbf{M}_0$, **1**, **2**, **3**, **4**, $\mathbf{m}_0$, $\mathbf{m}_1$, $\mathbf{m}_2$, $\mathbf{m}_3$, $\mathbf{m}_4$ are 2D homogeneous coordinates. This equation means that the cross ratio of the pencil of lines $\mathbf{M}_0$ (**1**, **2**, **3**, **4**) (the pencil of lines with $\mathbf{M}_0$ as the vertex and through **1**, **2**, **3**, **4**) is equal to the cross ratio of the pencil of lines

$\mathbf{m}_0$ ($\mathbf{m}_1$, $\mathbf{m}_2$, $\mathbf{m}_3$, $\mathbf{m}_4$).The reason is as follows. As shown in Fig. 2, the space plane A_1 intersects with the pencil of planes $\mathbf{OO}_c$ (**1**, **2**, **3**, **4**) at the pencil of lines $\mathbf{M}_0$ (**1**, **2**, **3**, **4**), and the image plane intersects with the pencil of planes $\mathbf{OO}_c$ (**1**, **2**, **3**, **4**) at the pencil of lines $\mathbf{m}_0$ ($\mathbf{m}_1$, $\mathbf{m}_2$, $\mathbf{m}_3$, $\mathbf{m}_4$). Therefore, the two cross ratios of $\mathbf{M}_0$ (**1**, **2**, **3**, **4**) and $\mathbf{m}_0$ ($\mathbf{m}_1$, $\mathbf{m}_2$, $\mathbf{m}_3$, $\mathbf{m}_4$) are both equal to the cross ratio of the pencil of planes $\mathbf{OO}_c$ (**1**, **2**, **3**, **4**). So, (1) is true.

The relationship (1) can be extended between perspective and omnidirectional image points as follows. Let $\mathbf{m}'_0$, $\mathbf{m}'_1$, $\mathbf{m}'_2$, $\mathbf{m}'_3$, $\mathbf{m}'_4$ be the perspective image points of $\mathbf{M}_0$, **1**, **2**, **3**, **4**, and **C** be the perspective camera center as in Fig. 2. The space plane A_1 intersects with the pencil of planes $\mathbf{CM}_0$ (**1**, **2**, **3**, **4**) at the pencil of lines $\mathbf{M}_0$ (**1**, **2**, **3**, **4**). And the perspective image plane intersects with the pencil of planes $\mathbf{CM}_0$ (**1**, **2**, **3**, **4**) at the pencil of lines $\mathbf{m}'_0$ ($\mathbf{m}'_1$, $\mathbf{m}'_2$, $\mathbf{m}'_3$, $\mathbf{m}'_4$). Therefore, the two cross ratios of $\mathbf{M}_0$ (**1**, **2**, **3**, **4**) and $\mathbf{m}'_0$ ($\mathbf{m}'_1$, $\mathbf{m}'_2$, $\mathbf{m}'_3$, $\mathbf{m}'_4$) are both equal to the cross ratio of the pencil of planes $\mathbf{CM}_0$ (**1**, **2**, **3**, **4**). Furthermore, from (1), we know the two cross ratios of $\mathbf{M}_0$ (**1**, **2**, **3**, **4**) and $\mathbf{m}_0$ ($\mathbf{m}_1$, $\mathbf{m}_2$, $\mathbf{m}_3$, $\mathbf{m}_4$) are equal. It follows that we have the geometric constraint:

$$\frac{[\mathbf{m}'_1\mathbf{m}'_3\mathbf{m}'_0][\mathbf{m}'_2\mathbf{m}'_4\mathbf{m}'_0]}{[\mathbf{m}'_2\mathbf{m}'_3\mathbf{m}'_0][\mathbf{m}'_1\mathbf{m}'_4\mathbf{m}'_0]} = \frac{[\mathbf{m}_1\mathbf{m}_3\mathbf{m}_0][\mathbf{m}_2\mathbf{m}_4\mathbf{m}_0]}{[\mathbf{m}_2\mathbf{m}_3\mathbf{m}_0][\mathbf{m}_1\mathbf{m}_4\mathbf{m}_0]}. \tag{2}$$

where $\mathbf{m}'_0$ varies with the space plane varying, of which locus is a line because different space planes intersect with the omnidirectional optical axis $\mathbf{OO}_c$ at different points. For example as shown in Fig. 2, A_2 is another space plane that intersects $\mathbf{OO}_c$ at $\mathbf{P}_0$. $\mathbf{P}_0$ is different from $\mathbf{M}_0$ and so their images under **C** as $\mathbf{p}'_0$ and $\mathbf{m}'_0$ are different.

3.2 Estimation of Geometric Constraint

In (2), if $\mathbf{m}_0$, $\mathbf{m}_1$, $\mathbf{m}_2$, $\mathbf{m}_3$,$\mathbf{m}'_0$, $\mathbf{m}'_1$, $\mathbf{m}'_2$, $\mathbf{m}'_3$ are known, and ($\mathbf{m}'_4$, $\mathbf{m}_4$) is a candidate match, then (2) forms a geometric constraint for match propagation. In other words, only those candidate matches which satisfy (2) may be stored in the result of the quasi-dense match and added to the list of seed points. We describe how to obtain $\mathbf{m}_0$, $\mathbf{m}_1$, $\mathbf{m}_2$, $\mathbf{m}_3$, $\mathbf{m}'_0$, $\mathbf{m}'_1$, $\mathbf{m}'_2$, $\mathbf{m}'_3$ in the following.

$\mathbf{m}_0$ is the principal point of the omnidirectional camera and it is very close to the image center or the contour center. Therefore, the center of the omnidirectional image or the contour can be used as an approximate solution of $\mathbf{m}_0$.

Different planar patches have different corresponding points of $\mathbf{m}_0$ in a perspective image. Therefore, it is necessary to compute the corresponding point of $\mathbf{m}_0$ in the perspective image for each planar patch. Let $\mathbf{m}'_0$ be the corresponding image point of $\mathbf{m}_0$ under a perspective view associated with a space plane. And let **S** be a set of matches from a space plane patch obtained by unconstrained propagation. **RANSAC** (Random Sample Consensus) [20] is used to estimate the corresponding point $\mathbf{m}'_0$ of $\mathbf{m}_0$. The detailed steps are as follows.

Step 1. Make a set $\mathbf{S}_1$ consisting of different quintuples of matches randomly selected from S. The quintuples are demoted as Q_k, k=1...N.

Step 2.Then, from each quintuple Q_k of $\mathbf{S}_1$, take the different quadruplets and denote one of the quadruplets as ($\mathbf{m}'_i$, $\mathbf{m}_i$), i = 1...4. Set up invariant equations on $\mathbf{m}'_0$ by (2):

$$[\mathbf{m}'_1\mathbf{m}'_3\mathbf{m}'_0][\mathbf{m}'_2\mathbf{m}'_4\mathbf{m}'_0][\mathbf{m}_2\mathbf{m}_3\mathbf{m}_0][\mathbf{m}_1\mathbf{m}_4\mathbf{m}_0]$$
$$-[\mathbf{m}'_2\mathbf{m}'_3\mathbf{m}'_0][\mathbf{m}'_1\mathbf{m}'_4\mathbf{m}'_0][\mathbf{m}_1\mathbf{m}_3\mathbf{m}_0][\mathbf{m}_2\mathbf{m}_4\mathbf{m}_0] = 0,$$

where $\mathbf{m}'_0$ is unknown with two elements. Take other quadruplets in Q_k, we obtain other equations. By repeating this process for different quadruplets in Q_k, we set up the equations on $\mathbf{m}'_0$.

Step 3. Solve the equations in Step 2 for $\mathbf{m}'_0 = (x, y, 1)$ associated with Q_k and the solution is denoted as s_k. Substitute the solution s_k into the corresponding rational equations (2) for other matches of $\mathbf{S}$. If the difference of the two sides in (2) is smaller than a threshold, these matches are inliers of Q_k.

Step 4. The quadruplet Q_k with the largest number of inliers is chosen. The solution of $\mathbf{m}'_0 = (x, y, 1)$ is again solved from all the equations in Step 2 from all the inliers. There are more than two quadric equations on $\mathbf{m}'_0$ available. We construct a matrix, whose each row is the coefficients of x^2, xy, y^2, x, y, 1 of each equation. Then a robust SVD-like approach is used to obtain a more accurate solution for $\mathbf{m}'_0 = (x, y, 1)$.

Similarly, we can obtain the corresponding $\mathbf{m}'_0$ for other plane patches.

3.3 Local Affine Transformations during Propagation

Our initial matches consist of point matches $(\mathbf{x}_i, \mathbf{x}'_i)$(the centroids of the matched elliptical regions), affine transformation matrices $\mathbf{A}_i$ and angles α_i between corresponding points' orientations.

The affine transformation matrices of new matches are computed by the formula $\mathbf{A} = \mathbf{S}'^{0.5}\mathbf{R}\mathbf{S}^{-0.5}$ in [15], where $\mathbf{S}$ and $\mathbf{S}'$ are the windowed second moment matrices, and $\mathbf{R}$ is a rotational matrix determined from a pair of corresponding epipolar lines. In this paper, the problem we deal with is pairs of perspective and omnidirectional images without prior camera pose determination and distortion correction. The method in [15] cannot be used to compute the rotation matrix $\mathbf{R}$ any longer. Our method to compute $\mathbf{R}$ is as follows.

To determine $\mathbf{R}$, a pair of corresponding directions in the two images is needed. The corresponding directions are achieved by computing the orientations of the corresponding points, similar to the way of assigning orientations to key-points in **SIFT** descriptor [6]. In detail, assume that the current seed is $(\mathbf{x}, \mathbf{x}')$, the corresponding transformation estimate is $\mathbf{A}$, and the angle between the orientations of $\mathbf{x}$ and $\mathbf{x}'$ is α. $(\mathbf{u}, \mathbf{u}')$ is a new seed match. A $(2(W + N) + 1) \times (2(W + N) + 1)$ square patch centered on the current seed point is extracted from the image which locally has a lower resolution (the local magnification factor is given by $abs(det\mathbf{A})$), and the corresponding area in the other image is obtained as the affinely mapped result from the square patch of the first image. The neighborhoods of $\mathbf{u}$ and $\mathbf{u}'$ used in computing their orientations are the black framed windows centered on $\mathbf{u}$ and $\mathbf{u}'$shown in Fig. 3. For $\mathbf{u}$ or $\mathbf{u}'$, an orientation histogram is formed from the gradient orientations of pixels within the neighborhood. The local peaks in the histogram are detected, and

those orientations are recorded. Comparing angles between all pairs of recorded orientations of $\mathbf{u}$ and $\mathbf{u}'$ with α, the orientations corresponding to the angle which is the nearest to α for $\mathbf{u}$ and $\mathbf{u}'$ are chosen.

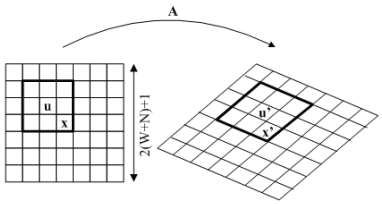

Fig. 3. The normalized image neighborhoods of size $(2(W+N)+1) \times (2(W+N)+1)$ for a seed point $(\mathbf{x}, \mathbf{x}')$. Here W=1 and N=2. The neighborhoods of $\mathbf{u}$ and $\mathbf{u}'$ used in computing their orientations are the black framed windows.

3.4 Propagation Algorithm

Combined with the above knowledge, the propagation algorithm between perspective and omnidirectional views is described briefly as follows.

1. Segment images into different plane patches. For each patch pair between perspective and omnidirectional views, do the following steps.

2. Perform initial matches: In order to get more initial matches, we can use a variety of detectors of which actions are complementary. Considering that ellipse equations of affine covariant regions of interest points are used to compute the affine transformation matrices for initial matches, we can choose those region detectors whose detected regions can be represented by ellipses, such as Hessian-Affine detector [10], Harris-Affine detector [8, 9], **MSER** (Maximally Stable Extremal Region) detector [7] and so on. In the experiments of Section 4, the Hessian-Affine detector and **MSER** detector are used to detect interest points in each view, and the detector regions are matched using the **SIFT** descriptor. For each initial match, affine transformation matrix $\mathbf{A}$ and angle α between corresponding points' orientations are estimated using the method in [15]. Because no suitable method for image matching between perspective and omnidirectional views is available at present, we borrow the method for perspective image matching. Therefore, the initial matches have a small number and an uneven distribution. This brings difficulties to the match propagation.

3. Make unconstrained propagation: From all the seed points by the strategy of Section 2 and by the method for computing affine transformations of Section 3.3, we obtain quasi-dense pixel correspondences.

4. Re-sample the quasi-dense pixel correspondences of step *3* by the method in [13, 14] to obtain the correspondences with even distribution; estimate the

corresponding point of the principle point by the method of Section 3.2 on the re-sampled points.

5. Make constrained propagation:The estimation in step *4* together with the equation (2) of Section 3.1 forms a geometric constraint. Constrained under this geometric relation, propagation from the same initial list of seeds by using a best-first strategy of Section 2 is performed.

6. Re-sample the obtained quasi-dense point correspondences of step *5*; re-estimate the corresponding point of the principle point with the re-sampled quasi-dense point correspondences by the method of Section 3.2.

7. Finally, remove outliers in the re-sampled quasi-dense point correspondences by the geometric constraint to obtain the final quasi-dense point correspondences with even distribution and high reliability.

4 Experiments

We conducted experiments with pairs of real images to test the algorithm described above. The perspective images are of size 640×480 pixels captured by a NIKON COOLPIX990 camera, and the omnidirectional images are of size 640×480 pixels captured by a NIKON FC-E8 fisheye converter mounted on a NIKON COOLPIX990 camera. The window size of **ZNCC** is 9×9. The neighborhood size from which new matches are searched is 7×7. An example is given below. In Fig. 4(a) there is a pair of perspective and omnidirectional images of size 640×480 pixels. Fig. 4(b) shows the initial matches obtained by the method in Section 3.4. There are 43 initial matches in all. Because there is no suitable method of image matching for perspective and omnidirectional views, we can see that the distribution of the initial matches is extremely uneven. The results of the step *3* to *7* of the algorithm in Section 3.4 are shown in Fig. 4(c)-4(g). It can be seen that the quasi-dense matches obtained by the algorithm described above are fairly dense and spread on most of the textured regions. The computation times per match for unconstrained and constrained propagations are 4.2 and 13.2 ms using our Matlab implementations.

To evaluate the correctness of the quasi-dense matches, we compute the corresponding points of the principle point with even distributed matches obtained manually.By section 3.1, we know (2) is the cross ratio equality on two pencils of lines, where $\mathbf{m}'_0$, $\mathbf{m}_0$ are respectively the vertexes. The line connecting $\mathbf{m}'_0$ and $\mathbf{m}'_i$ is the line in the pencil of $\mathbf{m}'_i$, and similarly the line $\mathbf{m}_0\mathbf{m}_i$ is that of $\mathbf{m}_i$. For each match ($\mathbf{m}'_i$, $\mathbf{m}_i$), the two lines $\mathbf{l}' \approx \mathbf{m}'_0\mathbf{m}'_i$ and $\mathbf{l} \approx \mathbf{m}_0\mathbf{m}_i$ are a pair of corresponding RP (Radiar and Polar) lines. We compute the corresponding RP line $\mathbf{l}$ in the omnidirectional image for $\mathbf{m}_i$ and the corresponding RP line $\mathbf{l}'$ in the perspective image for $\mathbf{m}'_i$ by (2) from different matches. The distances between $\mathbf{m}_i$ and $\mathbf{l}$, $\mathbf{m}'_i$ and $\mathbf{l}'$ are computed as the error of the match ($\mathbf{m}'_i$, $\mathbf{m}_i$). The match error map of the constrained propagation is shown in Fig. 4(h). The resolution of the perspective image is higher than the one of the omnidirectional image. Therefore, there are many points in the perspective image that have no corresponding points in the omnidirectional image. So the black points (with

very small match errors) in the match error map of the perspective image on the left of Fig. 4(h) are fairly sparse and there are many white points (that have no corresponding points in the omnidirectional image) evenly distributed in the maps. However, we can see from the figures that most of the matches are right.

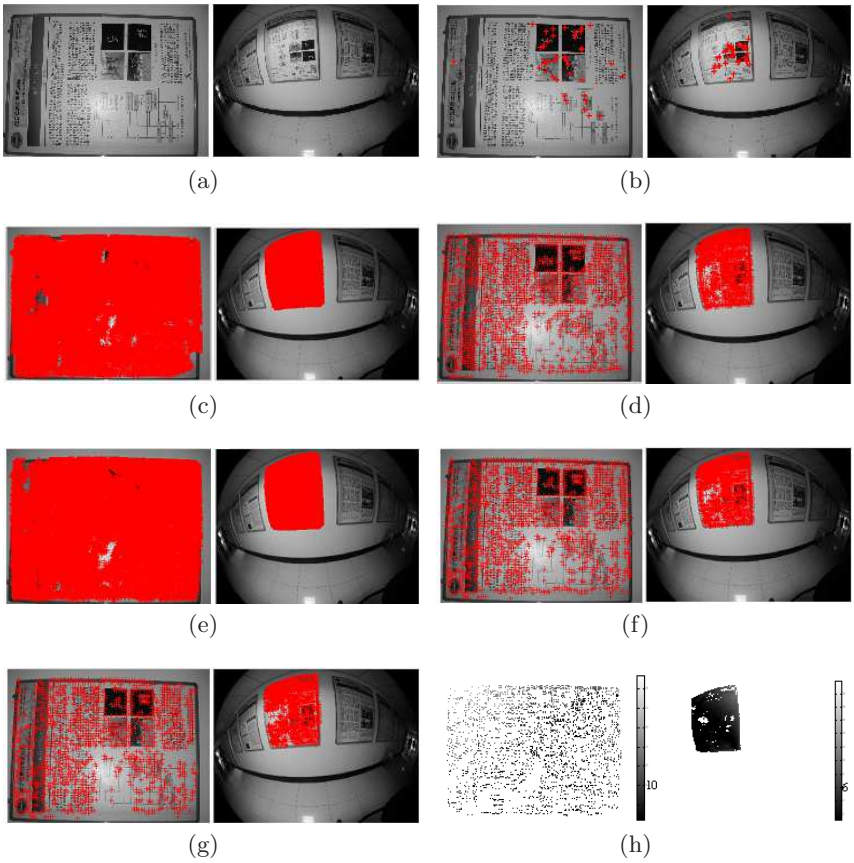

Fig. 4. (a)The perspective and omnidirectional images.(b)43 initial matches.(c)33977 matches obtained by unconstrained propagation.(d)1573 matches obtained by re-sampling.(e)32658 matches obtained by constrained propagation.(f)1506 matches obtained by re-sampling again.(g)1410 final quasi-dense matches.(h)The match error map of constrained propagation.

In the following experiment, the object in the omnidirectional image is close to the edge of the image and has severe distortions. The sizes of the images are also 640×480 pixels. The perspective and omnidirectional images used and the initial matches are shown in Fig. 5(a) and 5(b). The results of the step *4* and *7* of the propagation algorithm described in Section 3.4 and the match error map of

constrained propagation are shown in Fig. 5(c)-5(f). It can be seen from Fig. 5(b) that there are false matches in the initial matches. However, the false matches in the initial matches don't spread to the quasi-dense matches as shown in Fig. 5(d). This indicates that the algorithm is robust to some extent. The percentage of error matches in Fig. 5(c) is 10.10638% and the one in Fig. 5(d) is 5.36193%. The geometric constraint reduced the percentage of error matches. The computation times per match for unconstrained and constrained propagations are 3.2 and 8.3 ms using our Matlab implementations.

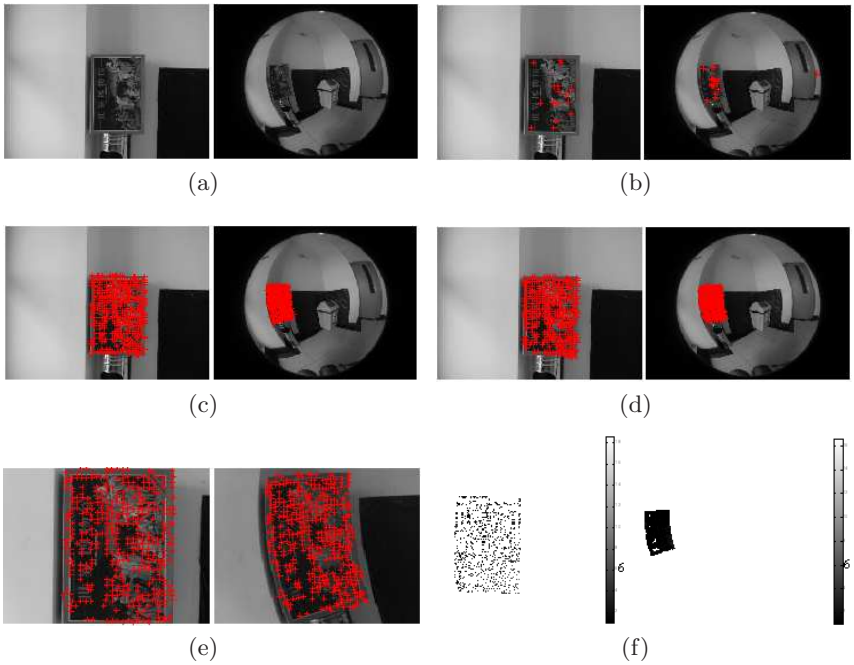

Fig. 5. (a)The perspective and omnidirectional images.(b)17 initial matches.(c)376 matches obtained by the step *4* of Section 3.4.(d)373 final quasi-dense matches.(e)The close-up of Fig. 5(d).(f)The match error map of constrained propagation.

The last experiment was conducted with image pairs of the multi-plane scene. The sizes of the images are also 640×480 pixels. The perspective and omnidirectional images used and the initial matches are shown in Fig. 6(a) and 6(b). The results of the step 7 of the propagation algorithm described in Section 3.4 and the match error map of constrained propagation are shown in Fig. 6(c) and 6(d). The percentages of error matches of the step *4* and *7* of the propagation algorithm described in Section 3.4 are respectively 7.87623% and 3.72671%. We can see that there are some false matches on the right of Fig. 6(d) (in the red circle), but the match errors of them are not big. The reason is as follows: this

region is lack of texture and the false match points lie very closely to the right corresponding lines, which make the **ZNCC** value and the geometric constraint ineffective. However, the false matches are removed by re-sampling, as shown in Fig. 6(c).

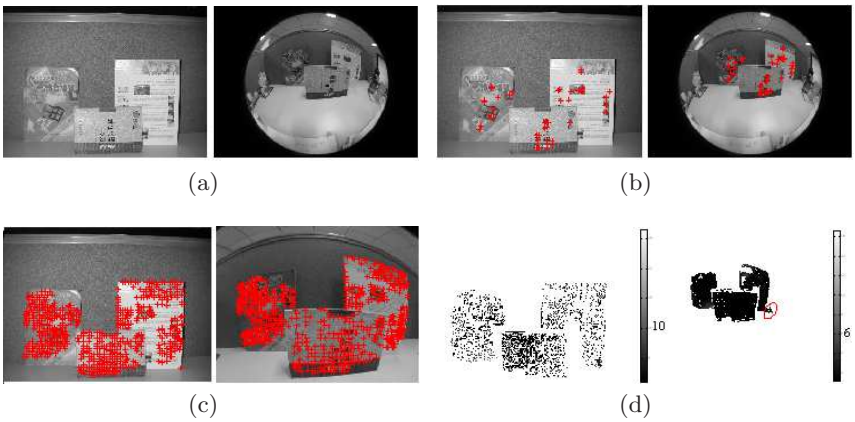

Fig. 6. (a)The perspective and omnidirectional images.(b)43 initial matches.(c)644 final quasi-dense matches (the right picture is magnified).(d)Match error map of constrained propagation.

5 Conclusions

We have proposed a match propagation algorithm between perspective and omnidirectional images of same scene consisting of planes, by which we can obtain quasi-dense point correspondences without calibration and distortion correction.

This proposed algorithm is developed from match propagation algorithms between perspective images. We use a linear transformation as an estimate for the local non-linear geometric transformation between perspective and omnidirectional image patches. A geometric invariant between perspective and omnidirectional views is established for the match propagation. Experiments have demonstrated that this algorithm can give good results of quasi-dense matching. In the future, we intend to investigate reconstructing the 3D scenes using the quasi-dense point correspondences obtained by this proposed algorithm.

Acknowledgement

This work was supported by the National Natural Science Foundation of China under grant No. 60773039, 60723005.

References

1. Ude, A., Gaskett, C., Cheng, G.: Foveated vision systems with two cameras per eye. In: Proceedings of the 2006 IEEE International Conference on Robotics and Automation. (2006) 3457–3462
2. Scassellati, B.: A binocular, foveated active vision system. (1999)
3. Atkeson, C.G., Hale, J.G., Pollick, F., Riley, M., Kotosaka, S., Schaal, S., Shibata, T., Tevatia, G., Ude, A., Vijayakumar, S., Kawato, M.: Using humanoid robots to study human behavior. IEEE Intelligent Systems **15** (2000) 46–56
4. Shibata, T., Vijayakumar, S., Conradt, J.J., Schaal, S.: Biomimetic oculomotor control. Adaptive Behavior **9** (2001) 189–208
5. Kozima, H., Yano, H.: A robot that learns to communicate with human caregivers. In: Proceedings of International Workshop on Epigenetic Robotics, Lund, Sweden (2001)
6. Lowe, D.: Distinctive image features from scale invariant keypoints. International Journal of Computer Vision **60** (2004) 91–110
7. Matas, J., Chum, O., Urban, M., Pajdla, T.: Robust wide-baseline stereo from maximally stable extremal regions. Image and Vision Computing **22** (2004) 761–767
8. Mikolajczyk, K., Schmid, C.: An affine invariant interest point detector. In: Proceedings of the 7th European Conference on Computer Vision. (2002) 128–142
9. Mikolajczyk, K., Schmid, C.: Scale & affine invariant interest point detectors. International Journal of Computer Vision **60** (2004) 63–86
10. Mikolajczyk, K., Tuytelaars, T., Schmid, C., et al: A comparison of affine region detectors. International Journal of Computer Vision **65** (2005) 43–72
11. Svoboda, T., Pajdla, T.: Matching in catadioptric images with appropriate windows, and outliers removal. In: Proceedings of the 9th International Conference on Computer Analysis of Images and Patterns, Berlin, German (2001) 733–740
12. Pohl, Z.: Omnidirectional vision–searching for correspondences. Master's thesis, Prague, Czech Republic (2001)
13. Lhuillier, M., Quan, L.: Match propagation for image-based modeling and rendering. IEEE Transactions on Pattern Analysis and Machine Intelligence **24** (2002) 1140–1146
14. Lhuillier, M., Quan, L.: A quasi-dense approach to surface reconstruction from uncalibrated images. IEEE Transactions on Pattern Analysis and Machine Intelligence **27** (2005) 418–433
15. Kannala, J., Brandt, S.S.: Quasi-dense wide baseline matching using match propagation. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, MN, USA (2007) 1–8
16. Faugeras, O.: Stratification of three-dimensional vision: Projective, affine and metric representations. Journal of the Optical Society of America **12** (1995) 465–484
17. Sturmfels, B.: Algorithms in Invariant Theory. Springer, New York (1993)
18. Wu, Y.H., Hu, Z.Y.: Geometric invariants and applications under catadioptric camera model. In: Proceedings of the Tenth IEEE International Conference on Computer Vision. (2005) 1547–1554
19. Geyer, C., Daniilidis, K.: Catadioptric projective geometry. International Journal of Computer Vision **45** (2001) 223–243
20. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM **24** (1981) 381–395