A CONSTRAINT TO IMPROVE THE RELIABILITY OF STEREO MATCHING USING THE RANK TRANSFORM

Jasmine Banks[1,2], *Mohammed Bennamoun*[1], *Kurt Kubik*[1] *and Peter Corke*[2,3]

[1]Space Centre for Satellite Navigation
Queensland University of Technology
Brisbane, Australia.
[2]Cooperative Research Centre for Mining Technology and Equipment
[3]CSIRO Manufacturing Science and Technology
{j.banks, m.bennamoun, k.kubik}@qut.edu.au, pic@cat.csiro.au

ABSTRACT

The rank transform is a non-parametric technique which has been recently proposed for the stereo matching problem. The motivation behind its application to the matching problem is its invariance to certain types of image distortion and noise, as well as its amenability to real-time implementation. This paper derives an analytic expression for the process of matching using the rank transform, and then goes on to derive one constraint which must be satisfied for a correct match. This has been dubbed the *rank order constraint* or simply the *rank constraint*. Experimental work has shown that this constraint is capable of resolving ambiguous matches, thereby improving matching reliability. This constraint was incorporated into a new algorithm for matching using the rank transform. This modified algorithm resulted in an increased proportion of correct matches, for all test imagery used.

1. INTRODUCTION

A fundamental problem faced by stereo vision algorithms is that of determining correspondences between two images which comprise a stereo pair. One non-parametric transform which has recently been proposed for the stereo matching problem is the rank transform[9]. The advantages of this transform include its invariance to radiometric distortion[2], and its amenability to fast hardware implementation.

Constraints such as the left–right consistency criterion and removal of locally anomalous disparities have been widely used by matching algorithms in order to identify and remove invalid matches[6, 7]. Although powerful, most of these constraints have little theoretical basis. In this paper, we theoretically derive and test a new constraint for matching using the rank transform. The process of matching using the rank transform is outlined in Section 2. Section 3 derives a constraint which must be satisfied for a correct match.

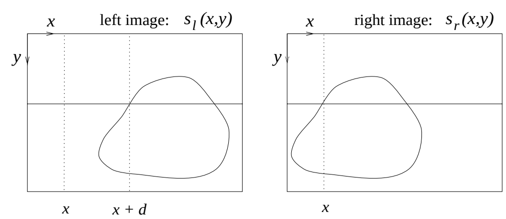


Figure 1: Stereo pair to be matched. The left and right images are denoted $s_l(x, y)$ and $s_r(x, y)$ respectively.

As described in Section 4, this constraint was tested using a number of stereo pairs. Finally, the main contributions of this work are summarised in Section 5.

2. MATCHING USING THE RANK TRANSFORM

2.1. Image Representation

As shown in Figure 1, the left and right images which comprise the stereo pair are denoted $s_l(x, y)$ and $s_r(x, y)$ respectively. Assuming the images are epipolar[4], a simple model for the relationship between corresponding image points is given by [8]:

$$s_r(x, y) = As_l(ax + d, y) + B + N(x, y) \tag{1}$$

where A and B are the contrast and brightness factors for radiometric distortion, and N represents noise. The terms a and d represent geometric distortion, and in particular, d is the disparity difference we wish to find.

2.2. The Matching Process

The matching process accepts an epipolar aligned stereo pair as input, and produces a *disparity map* as output. Figure 2 illustrates the process. This section describes each step of this process in detail.

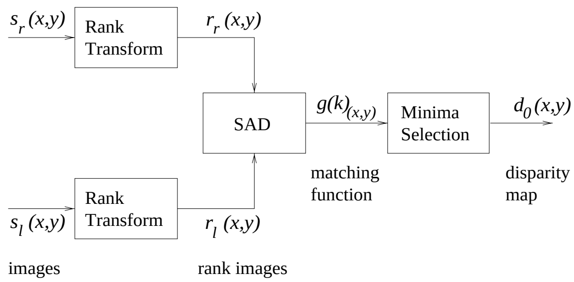


Figure 2: The matching process using the rank transform.

2.2.1. *The Rank Transform*

The rank transformation process involves passing the rank window over the image, and at each point, counting the number of pixels in the rank window whose value is less than the centre pixel. The rank transform may be expressed as:

$$r(x,y) = R - \sum_{(i,j)\in W} U\big[s(x+i,y+j) - s(x,y)\big] \quad (2)$$

where $U[t]$ is the unit step function, R is the number of pixels in the rank window and $(i,j) \in W$ indicates the neighbourhood of the rank window.

2.2.2. *Computation of SAD Matching Metric*

The Sum of Absolute Differences (SAD) matching metric[1] provides a measure of the *similarity* between pixel regions. Given a *template* window, centred on $r_r(x,y)$, the SAD metric is computed for a series of *candidate* windows, centred on $r_l(x+k,y)$, where the test disparity k is varied in integer increments from 0 to d_{max}. This series of SAD scores is referred to as a *match function*, and is computed as follows

$$\begin{aligned} g(k)_{(x,y)} &= \mathrm{SAD}\big(r_r(x,y), r_l(x+k,y)\big) \\ &= \sum_{(m,n)\in M} \big|r_r(x+m,y+n) - r_l(x+k+m,y+n)\big| \end{aligned} \quad (3)$$

where $(m,n) \in M$ indicates summation of over the match window. An example of a match function is shown in Figure 4(a).

2.2.3. *Selection of Minima*

This involves selection of the disparity at which the match function is a minimum, which is the disparity at which the template and candidate windows are most similar. The selection of the minima is expressed as

$$\begin{aligned} d_0(x,y) = k \mid \Big[&\big(\forall p\,(p<k)\; g(k)_{(x,y)} < g(p)_{(x,y)}\big) \\ &\text{AND } \big(\forall p\,(p>k)\; g(k)_{(x,y)} \le g(p)_{(x,y)}\big)\Big] \end{aligned} \quad (4)$$

The disparity values $d_0(x,y)$ together comprise a *disparity map*. It is desirable that these disparity values correspond to the d term in Equation (1), the true difference in location of the pixel patterns in $s_l(x,y)$ and $s_r(x,y)$.

3. A MATCHING CONSTRAINT

This section formulates an analytic expression for the match function and uses this to derive one possible constraint for a correct match. The rank transformed images are computed from Equation (2) as follows

$$r_l(x,y) = R - \sum_{(i,j)\in W} U\big[s_l(x+i,y+j) - s_l(x,y)\big] \quad (5)$$

$$r_r(x,y) = R - \sum_{(i,j)\in W} U\big[s_r(x+i,y+j) - s_r(x,y)\big] \quad (6)$$

Substituting Equations (5) and (6) into the match function of Equation (3) results in

$$g(k)_{(x,y)} = \sum_{(m,n)\in M} \left| \sum_{(i,j)\in W} U(D_l) - U(D_r) \right| \quad (7)$$

where

$$\begin{aligned} D_l &= s_l(x+k+m+i, y+n+j) \\ &\quad - s_l(x+k+m, y+n) \end{aligned} \quad (8)$$

$$D_r = s_r(x+m+i, y+n+j) - s_r(x+m, y+n) \quad (9)$$

The optimum disparity is selected as the disparity k at which Equation (3) is a minimum. As shown in [3], differentiating Equation (3) with respect to the test disparity, k results in

$$\begin{aligned} g'(k)_{(x,y)} = \sum_{(m,n)\in M} \Bigg\{ & sgn\Big(\sum_{(i,j)\in W} U(D_l) - U(D_r) \Big) \\ & \Big(\sum_{(i,j)\in W} \delta(D_l)(D_l') \Big) \Bigg\} \end{aligned} \quad (10)$$

where D_l' is given by

$$\begin{aligned} D_l' &= s_l'(x+k+m+i, y+n+j) \\ &\quad - s_l'(x+k+m, y+n) \end{aligned} \quad (11)$$

The first derivative will be zero if the sgn term is zero for all (m,n,i,j). Since the only possible values for the function $U[x]$ are 0 or 1, the conditions for the sgn term to be zero are

$$\begin{matrix} U[D_l] = 0 & \text{or} & U[D_l] = 1 \\ U[D_r] = 0 & & U[D_r] = 1 \end{matrix} \quad (12)$$

for all (m,n,i,j). Substituting Equations (8) and (9) into (12) yields

$$\begin{aligned} &s_l(x+k+m, y+n) > s_l(x+k+m+i, y+n+j), \\ &s_r(x+m, y+n) > s_r(x+m+i, y+n+j) \end{aligned} \quad (13)$$

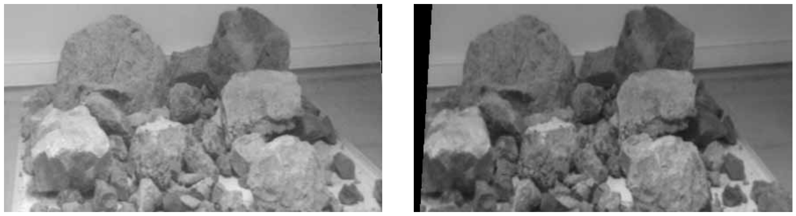


Figure 3: Stereo pair of rocks[5].

$$s_l(x+k+m+i, y+n+j) \geq s_l(x+k+m, y+n),$$
$$s_r(x+m+i, y+n+j) \geq s_r(x+m, y+n) \qquad (14)$$

Equations (13) and (14) together form one constraint for a correct match. A more detailed derivation and analysis of this constraint will be forthcoming in [3]. The constraint depends on the relative ordering of pixels in rank windows centred on every pixel of the template and candidate windows. Consequently, it is referred to as the *rank order constraint* or simply the *rank constraint*.

4. TESTING AND RESULTS

In order to compute a measure of how well the rank constraint is satisfied, the pixel values in the template and candidate windows are examined, and a score incremented each time the constraint is *not satisfied*. This measure score may be computed for each test disparity, resulting in a *constraint evaluation function*.

Initial tests were carried out using a contrived stereo pair consisting of two images displaced by a known amount. Since the image is matched with a displaced version of itself, the true disparity, d is precisely known, and furthermore, there is no noise or radiometric or geometric distortion, i.e., $N(x,y) = 0$, $a = 1$, $A = 1$ and $B = 0$ in Equation (1). The constraint was always completely satisfied at the correct disparity, thus confirming the validity of the constraint[3].

Figure 4 illustrates the ability of the rank constraint to resolve ambiguous matches. A match function and constraint evaluation function, derived from the stereo pair of Figure 3, are shown. For a good match, the match function should have a single dominant minima. However, Figure 4 has two main minima, at disparities of 3 and 23, where 23 is in fact the correct disparity. Unfortunately, the minima at a disparity of 3 is slightly less than the one at 23, which would result in an incorrect disparity being returned. However, the constraint evaluation function has one dominant minima at a disparity of 23, and is thus able to be used for resolving the ambiguous match. Note that the constraint function does not reach zero at the correct disparity, due to the presence of real image distortions. Further testing has shown that the rank constraint is able to resolve a large number of cases of ambiguous matches, not only for the stereo pair of Figure 3, but also for other test pairs used.

The rank constraint has been incorporated into the minima selection stage of the matching algorithm of Section 2. Selection of the minima proceeds as follows

for each minima of the match function
    compute the constraint score
select the minima with the optimum constraint score

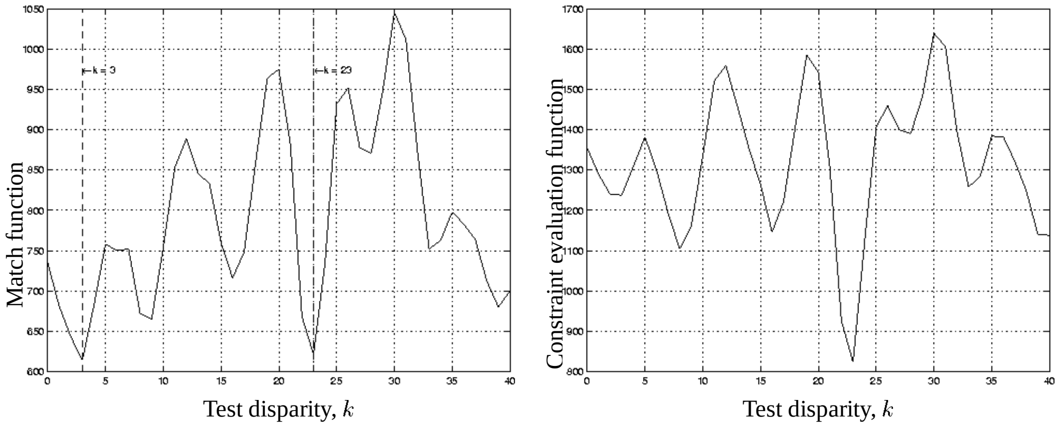
The modified algorithm was implemented and tested using a number of test stereo pairs. After matching was carried out, some well known techniques for removing invalid matches were applied to the resulting disparity maps, including left–right consistency checking, removal of locally anomalous matches and removal of matches for bland areas[6, 7]. The results for the stereo pair of Figure 3 are shown in Figure 5. From a visual inspection of the disparity maps, it can be seen that the modified algorithm using the constraint has correctly matched some areas which were not matched using the original rank algorithm. Most of these areas corresponded to ambiguous matches, which again illustrates the ability of the rank constraint to resolve ambiguous matches.

5. CONCLUSION

The contributions arising from this work include the derivation of one constraint which must be satisfied for a correct match, namely the *rank order constraint*. Experimental results show that this constraint is capable of resolving ambiguous matches. A novel matching algorithm incorporating this constraint has also been proposed. Testing performed using a number of stereo pairs have shown that the modified algorithm consistently resulted in a increased proportion of correct matches, thereby improving matching reliability.

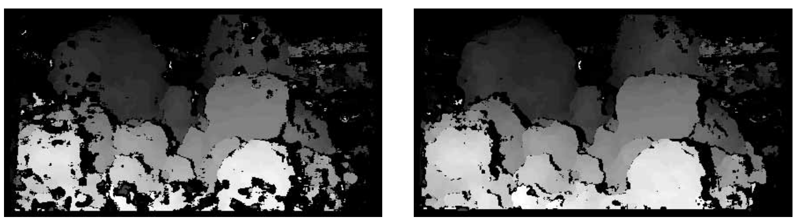
6. REFERENCES

[1] P. Aschwanden and W. Guggenbühl. Experimental results from a comparative study on correlation-type reg-



(a) (b)

Figure 4: (a) match function and (b) constraint evaluation function, derived from the stereo pair of Figure 3, using a template match window centred on $(188, 151)$. In each case, a rank window size of 5×5 and a match window size of 11×11 were used. The match function has two main minima, at disparities of 3 and 23, where 23 is in fact the correct disparity. Unfortunately, the minima at a disparity of 3 is slightly less than the one at 23, which would result in an incorrect disparity being returned. However, constraint function has one dominant minima at a disparity of 23, and is thus able to resolve the ambiguous match.



(a) (b)

Figure 5: Disparity results obtained for the stereo pair of Figure 3, using (a) the original rank matching algorithm and (b) the modified algorithm using rank constraint. In each case a rank window of 5×5 and a match window of 11×11 were used. In these images, black areas indicate invalid matches which were removed, while grey-scale areas indicate the remaining disparity values. From a visual inspection of the disparity maps, it can be seen that some areas which were not matched in (a) were correctly matched in (b).

istration algorithms. In *Robust Computer Vision*, pages 268–289. Wickmann, 1993.

[2] J. Banks, M. Bennamoun, and P. Corke. Fast and robust stereo matching algorithms for mining automation. In *Proc. IAIF'97*, Nov 1997.

[3] J. Banks, M. Bennamoun, P. Corke, and K. Kubik. Analysis of the rank transform for stereo matching. *In preparation*.

[4] S. Barnard and M. Fischler. Computational stereo. *Computing Surveys*, 14(4):553–572, Dec 1982.

[5] R. Bolles, H. Baker, and M. Hannah. The JISCT stereo evaluation. In *Image Understanding Workshop*, pages 263–274. DARPA, 1993.

[6] S. Cochran and G. Medioni. 3-D surface description from binocular stereo. *IEEE Trans. on PAMI*, 14(10):981–994, Oct 1992.

[7] P. Fua. A parallel stereo algorithm that produces dense depth maps and preserves image features. *Machine Vision and Applications*, 6:35–49, 1993.

[8] C. Tomasi and R. Manduchi. Stereo matching as a nearest neighbour problem. *IEEE Trans. on PAMI*, 20(3):333–340, Mar 1998.

[9] R. Zabih and J. Woodfill. Non-parametric local transforms for computing visual correspondence. In *Proc. ECCV'94*, 1994.