A STEREO LINE MATCHING TECHNIQUE FOR AERIAL IMAGES
BASED ON A PAIR-WISE RELATION APPROACH

A. O. Ok a *, J. D. Wegner b, C. Heipke b, F. Rottensteiner b, U. Soergel b, V. Toprak a

a Dept. of Geodetic and Geographic Information Tech., Middle East Technical University, 06531 Ankara, Turkey
b Institute of Photogrammetry and Geoinformation, University of Hannover, 30167 Hannover, Germany

- (oozgun, toprak)@metu.edu.tr
- (wegner, heipke, rottensteiner, soergel)@ipi.uni-hannover.de

Commission I, WG I/4

KEY WORDS: pair-wise line matching, post-processing, line extraction, stereo aerial images

ABSTRACT:

In this study, we developed a new pair-wise relation based approach for the matching of line features from stereo aerial images. To solve the final matching inconsistencies, we propose an iterative pair based post-processing algorithm in which the matching inconsistencies are eliminated using three novel measures and a final similarity voting scheme. The approach is tested over four urban test sites with various built-up characteristics, and for all test sites, we achieved a stereo line matching performance of 98%. The overall results indicate that the proposed approach is highly robust for the line features extracted in (very) dense urban areas.

1. INTRODUCTION

Corresponding lines in overlapping aerial images can be used for different purposes such as 3D object extraction, improving the automated triangulation, image registration, motion analysis etc. However, line matching in ultra high resolution (6–8 cm) stereo aerial images is a very challenging task due to various reasons; substantial change in viewpoints, inconsistency of line endpoint locations, the limitations of the geometric constraints imposed, lack of rich textures in line local neighbourhood, repetitive patterns etc. Up to now, a significant number of research papers have been published in this field; however, in a stereo environment, the ambiguity problem of line matching is an issue that remain unsolved. The major problem in line matching arises from the lack of measure(s) and/or constraint(s) for line features that are invariant under different viewing conditions. Existing geometric attributes for line matching in the stereo geometry is strictly limited. For example, Zhang (2005) mentioned the major problems of the available geometric constraints for line features, and finally, utilized only the orientation of the line segments as a single geometric constraint. The information around the line local neighbourhood is also well issued by most of the researchers; for example, Schmid and Zisserman (1997) proposed a direct and warped correlation measures computed around the line local neighborhoods observed by short and long range motions, respectively. The color and chromatic information within the local neighbourhood was also well issued in elsewhere (Scholze et. al., 2000; Zhang and Baltsavias, 2000, Herbert et. al., 2005). Recently, Wang et. al. (2009) proposed a new measure that takes into account the information based on the gradient orientation around the line local neighbourhoods and presented some good results for a number of close range datasets. However, all those measures are almost non-discriminative by their own for the aerial image case and suffer from the same problem, repetitive patterns, where the information extracted from local neighbourhoods of very different lines has similar information (Fig. 1). Therefore, for example, the strength of the work of Schmid and Zisserman (1997) relies to the post-processing stage where the epipolar ordering constraint is forced over line features for the disambiguation. However, the ordering constraint has some critical drawbacks as well; (i) some certain lines especially belonging to thin objects (mostly details of buildings) will be unquestionable lost, (ii) the matching failures especially in occluded areas that do not violate the ordering constraint could not be detected and eliminated. To be specific, almost all the previous work related to line matching relies on various descriptors specialized for one to one line matching in which the relations between the line features are not taken into account. However, the integration of those line relations during matching expose new constraints and further possibilities to improve the quality and the performance of the line matching.

2. METHODOLOGY

2.1 Line Extraction and Pair-wise Line Matching

In a recent work, we have presented a novel approach for the pair-wise matching of line features (Ok et. al., 2010). In this paper, we only briefly summarize the algorithm and refer the reader to the reference for further details. The algorithm consists of two main steps; (i) straight line extraction and (ii) the stereo matching of the extracted lines with a pair-wise approach. During the first step, in order to maximize the performance of the line detection, existing multispectral information in aerial images was fully utilized throughout the steps of pre-processing and edge detection. To accurately describe the straight edge segments, a principal component analysis technique was adapted and the extracted segments were converted to their line counterparts using an iterative Ransac algorithm. To establish the pair-wise line correspondences between the stereo images,

Figure 1. Two examples of line segments commonly observed in repetitive patterns and their very similar local neighbourhoods.
a new pair-wise stereo line matching approach was developed. For each pair in the left image, the best candidate line pair in the right image was assigned after a weighted pair-wise matching similarity score which was computed over a total of eight measures; an epipolar, three geometric, two photometric, a correlation and a spatiogram constraint. All the measures were normalized from 0 to 1 prior to the calculation, and the total similarity result was computed as the average of all similarities.

The main problem of the pair-wise matching presented is that it does not always guarantee one to one matches for each line. Based on our experiences, after the pair-wise matching, the ambiguities mostly occur for the lines that are adjacent located within a very short perpendicular distance. A typical example is given in Fig. 2. This is mainly due to two explicit reasons; (i) the lines that are very close to each other that belong to the same object (building, road etc.) reveal similar pair-wise characteristics and (ii) since we apply relaxed thresholds during pair-wise matching (especially for the epipolar intersection), very close lines are mostly susceptible to satisfy those thresholds. Therefore, in this paper, a great care has been devoted to the post-processing stage and a new iterative disambiguation algorithm is developed. For this purpose, we combined three novel measures during the selection of the best line correspondences. (i) The first measure relies on the gradient orientation information in the local neighbourhood of lines which is computed using a recently proposed dense matching measure, Daisy (Tola et. al., 2010). Since the original Daisy measure is point based, in this study, the measure is extended and adapted to fulfil the requirements of the linear features and their local neighbourhood. (ii) The second measure, the Redundancy, is computed from the entire pair-wise matches based on the fact that a single line is allowed to have a part in different pair combinations. Thus, after the pair-wise matching, there is a quite large number of matching redundancy available for most of the line correspondences. By this way, the redundancy measure gives a possibility to understand and integrate a local matching support for lines during the disambiguation process. (iii) The third measure is computed from the results of each individual pair-wise matching. Since we assigned the best pair using a pair-wise matching similarity score, this information can also be utilized during the post-processing, since the quality of the pair matches inherently determined by the quality of the line correspondences in each pair. We integrated those three measures for the final disambiguation process in an exclusively developed iterative way, in which the matching inconsistencies are eliminated using nearest/next ratios and a final similarity voting scheme.

2.2 Measures Utilized During Post-Processing

2.2.1 Daisy Measure

In recent years, the gradient orientation histograms has proven to be robust to distortions (up to a level) and found to be successful in terms of point matching when compared to the classical pixel-based measures such as cross-correlation and pixel differencing. Some good examples can be found in (Lowe, 2004; Mikolajczyk and Schmid, 2005; Bay et. al., 2006). More recently, Tola et al. (2010) proposed a dense matcher, Daisy, which is also proven to be much more efficient during the computation of the gradient orientation histograms. In the line matching context, up to our knowledge, the only study that takes into account the gradient orientation around the line local neighbourhood was proposed by Wang et al. (2009). However, as we already mentioned in the first section, for aerial images, the final decisions of the line matching that are only based on the information obtained from the line local neighbourhoods could be ambiguous. However, the information within those neighbourhoods may reveal some hints and may provide opportunities to indicate and eliminate the indisputably wrong matches.

In this study, we selected the Daisy as a fundamental local neighbourhood measure for the post-processing due to two explicit reasons; (i) great efficiency and speed during the computation of the gradient orientation histograms, (ii) its circular, symmetric shape and isotropic kernel structure turns out in a small overhead during the computation of the measure for different line orientations. Here, first, we only briefly review the original point-based Daisy measure and refer the reader to the reference for further details. Thereafter, we will introduce new adaptations for the Daisy and present how efficiently the measure could be utilized for capturing the line local neighbourhoods.

The Daisy descriptor is given in Fig. 3 (Tola et. al., 2010). In the descriptor, each circle represents a region where the radius is proportional to the standard deviations of the Gaussian kernels and the “+” sign represents the pixel locations of the convolved orientation map centers where the descriptor is computed. Daisy is controlled by a total of 4 parameters; where $R$ is the distance from the center pixel to the outer most grid point, $Q$ is the number of convolved orientation levels, $T$ is the number of histograms at a single layer, and $H$ is the number of bins in the histogram. For a given input image, first, depending on the number of bins $H$, orientation maps are computed. Each orientation map is then incrementally convolved with Gaussian kernels of different sigma values to obtain convolved orientation maps. At each pixel location illustrated in Fig. 3, a vector made of values from the convolved orientation maps are computed. Let $h_u(u_v)$ represent the vector made of the values at location $(u_v)$ in the orientation maps after convolution by a Gaussian kernel of standard deviation $\Sigma$; and let $Q$ represents the number of different circular layers, then the Daisy descriptor $D(u_v,v_0)$ for location $(u_v,v_0)$ is defined as:

![Figure 2. The matching ambiguities after pair-wise matching.](image-url)
The Daisy descriptor

\[ D(u_0, \theta_0) = \left[ h_{1,0}^X (u_0, \theta_0), h_{1,0}^Y (u_0, \theta_0), \ldots, h_{K,0}^X (u_0, \theta_0), h_{K,0}^Y (u_0, \theta_0) \right]^T \]

where \( I(u, v, R) \) is the location with distance \( R \) from \((u, v)\) in the direction given by \( j \) when the directions are quantized into the \( T \) values (Tola et al., 2010).

Since the Daisy descriptor is point-based (it belongs to the center grid point), in this study, the measure is extended and adapted to fulfil the requirements of the line features and their local neighbourhood. First, we centralize the center grid point of the descriptor to the center of the overlapping parts of the line segments which are defined by point to point correspondence (Ok et al., 2010). Next, to achieve rotation invariance over gradient vectors, we rotate the Daisy grid and align the direction vector (Fig. 3) of the descriptor with the orientation of each line. Since the amount of rotation must be adjusted for all lines based on their angle values in image space, during this procedure, we fully utilize one of the main advantages of the Daisy in which we only circularly shift the final orientation histograms to compute the descriptor. To achieve invariance to perspective distortion exactly on the line segments, for each line, we utilize adaptive \( R \) values for the Daisy grid (distance from the center pixel to the outer most grid point). The original Daisy measure has a specific constant \( R \) value; however, adaptive \( R \) values for line segments could be computed with the knowledge of the overlapping parts after imposing point to point correspondence. Since we apply this correspondence during the initial pair-wise matching, it does not bring any further overhead during the computation of the measure. In addition, it is apparent that we don’t have any knowledge about the surfaces attached to the lines in their neighbourhoods; thus, we further utilize the adaptively computed \( R \) values for entire Daisy grid points. After these adaptations, for the computation of the similarities, we divide the Daisy grid points into two separate classes and produce two constant grid binary masks \( \{ M_+ (x) \} \) for each class; the grid points that are located (i) above the line, and (ii) below the line. Thus, we perform the similarity computations independently for each grid class. Moreover, we also mask out the vector made values, \( h_{f(X,Y)} \), from the descriptor matrix \( D(u_0, \theta_0) \) whose grid locations are exactly on the line. This is due to the reason that if one of the sides of the lines is occluded, then the histograms computed for the points that are exactly on the lines have no reason to resemble each other. Therefore, we exclude those pixel locations and their histograms from the Daisy measure. For the computation of the dissimilarities between two Daisy descriptors, Tola et al. (2010) proposed a Euclidean difference metric:

\[ D = \frac{1}{\sum_{k=1}^{K} M[k]} \sum_{k=1}^{K} M[k] \left\| D_i[k] - D_j[k] \right\|_2 \]

where \( S \) is the number of grid points, \( M[k] \) is the \( k \)th element of the binary mask \( M \), and \( D_i[k] \) is the \( k \)th histogram \( h \) in \( D(x) \) computed image \( i \). However, we observed that, although the metric is successful in most of the cases, it completely ignores the cross-correlation between the two descriptors, \( D_i \) and \( D_j \). Thus, we define a modified similarity (\( M_S \)) metric that can be jointly utilized with the cross-correlation:

\[ M_S = \frac{1}{1 + \left( \sum_{k=1}^{K} M[k] \left\| D_i[k] - D_j[k] \right\|_2 \right)^2} \]

First, the normalization coefficient in Eq. 2 is not necessary any longer since our binary masks have constant number of points for each side of the lines. After the modification, the similarity metric produces values between 0 and 1, and in order to be more discriminative, we take the square of the total dissimilarity, thus, we further penalize the higher dissimilarities more discriminative, we take the square of the total dissimilarity, thus, we further penalize the higher dissimilarities more discriminative, we take the square of the total dissimilarity, thus, we further penalize the higher dissimilarities.

The final (\( \cdot \)) – operator in Eq. 5 ensures that the final Daisy similarity metric produces values between 0 and 1, and in order to be more discriminative, we take the square of the total dissimilarity, thus, we further penalize the higher dissimilarities more discriminative, we take the square of the total dissimilarity, thus, we further penalize the higher dissimilarities more discriminative.

Finally, since the similarities in Eq. 3 and 4 are computed independently for both sides of lines (for the above and below grid points), we propose our final Daisy similarity (\( \text{Sim}_D \)) for line matching as:

\[ \text{Sim}_D = \left( M^{\text{above}} + M^{\text{below}} \right)^+ \left( \mu^{\text{above}}, \mu^{\text{below}} \right)^+ \]

where \( \mu \) and \( s( \cdot ) \) operators denote the mean and standard deviations, respectively. Note that, in Eq. 4, similar to in Eq. 3, the correlation is also squared in order to give more weight to high similarity values, and to be even more discriminative.

Finally, since the similarities in Eq. 3 and 4 are computed independently for both sides of lines (for the above and below grid points), we propose our final Daisy similarity (\( \text{Sim}_D \)) for line matching as:

\[ \text{Sim}_D = \left( M^{\text{above}} + M^{\text{below}} \right)^+ \left( \mu^{\text{above}}, \mu^{\text{below}} \right)^+ \]

where \( \mu \) and \( s( \cdot ) \) operators denote the mean and standard deviations, respectively. Note that, in Eq. 4, similar to in Eq. 3, the correlation is also squared in order to give more weight to high similarity values, and to be even more discriminative.

Finally, since the similarities in Eq. 3 and 4 are computed independently for both sides of lines (for the above and below grid points), we propose our final Daisy similarity (\( \text{Sim}_D \)) for line matching as:

\[ \text{Sim}_D = \left( M^{\text{above}} + M^{\text{below}} \right)^+ \left( \mu^{\text{above}}, \mu^{\text{below}} \right)^+ \]

where \( \mu \) and \( s( \cdot ) \) operators denote the mean and standard deviations, respectively. Note that, in Eq. 4, similar to in Eq. 3, the correlation is also squared in order to give more weight to high similarity values, and to be even more discriminative.

Finally, since the similarities in Eq. 3 and 4 are computed independently for both sides of lines (for the above and below grid points), we propose our final Daisy similarity (\( \text{Sim}_D \)) for line matching as:

\[ \text{Sim}_D = \left( M^{\text{above}} + M^{\text{below}} \right)^+ \left( \mu^{\text{above}}, \mu^{\text{below}} \right)^+ \]

where \( \mu \) and \( s( \cdot ) \) operators denote the mean and standard deviations, respectively. Note that, in Eq. 4, similar to in Eq. 3, the correlation is also squared in order to give more weight to high similarity values, and to be even more discriminative.

Finally, since the similarities in Eq. 3 and 4 are computed independently for both sides of lines (for the above and below grid points), we propose our final Daisy similarity (\( \text{Sim}_D \)) for line matching as:

\[ \text{Sim}_D = \left( M^{\text{above}} + M^{\text{below}} \right)^+ \left( \mu^{\text{above}}, \mu^{\text{below}} \right)^+ \]

where \( \mu \) and \( s( \cdot ) \) operators denote the mean and standard deviations, respectively. Note that, in Eq. 4, similar to in Eq. 3, the correlation is also squared in order to give more weight to high similarity values, and to be even more discriminative.

Finally, since the similarities in Eq. 3 and 4 are computed independently for both sides of lines (for the above and below grid points), we propose our final Daisy similarity (\( \text{Sim}_D \)) for line matching as:

\[ \text{Sim}_D = \left( M^{\text{above}} + M^{\text{below}} \right)^+ \left( \mu^{\text{above}}, \mu^{\text{below}} \right)^+ \]

where \( \mu \) and \( s( \cdot ) \) operators denote the mean and standard deviations, respectively. Note that, in Eq. 4, similar to in Eq. 3, the correlation is also squared in order to give more weight to high similarity values, and to be even more discriminative.

Finally, since the similarities in Eq. 3 and 4 are computed independently for both sides of lines (for the above and below grid points), we propose our final Daisy similarity (\( \text{Sim}_D \)) for line matching as:

\[ \text{Sim}_D = \left( M^{\text{above}} + M^{\text{below}} \right)^+ \left( \mu^{\text{above}}, \mu^{\text{below}} \right)^+ \]

where \( \mu \) and \( s( \cdot ) \) operators denote the mean and standard deviations, respectively. Note that, in Eq. 4, similar to in Eq. 3, the correlation is also squared in order to give more weight to high similarity values, and to be even more discriminative.

Finally, since the similarities in Eq. 3 and 4 are computed independently for both sides of lines (for the above and below grid points), we propose our final Daisy similarity (\( \text{Sim}_D \)) for line matching as:

\[ \text{Sim}_D = \left( M^{\text{above}} + M^{\text{below}} \right)^+ \left( \mu^{\text{above}}, \mu^{\text{below}} \right)^+ \]

where \( \mu \) and \( s( \cdot ) \) operators denote the mean and standard deviations, respectively. Note that, in Eq. 4, similar to in Eq. 3, the correlation is also squared in order to give more weight to high similarity values, and to be even more discriminative.
Only one pair-wise relation indicates the segment #1, four out of five corresponds to the segment #a in Table 1a, among the total of five pair-wise relations that involve very limited chance to occur in multiple times. For example, in the ambiguities occur due to accidental alignments and has a uniqueness constraint must be handled carefully by taking into account the colinearity of the fragmented ones (segments #2 and #3). In this case, the uniqueness constraint is a part of the right image. To solve the ambiguity, we provide the new redundancy measure (SimR) for a line pair as:

\[ Sim_R = \frac{1}{N} \sum_{i=1}^{N} \frac{\min(j \mid d^{[i]}_j \neq 0 \text{ and } d^{[R]}_j \neq 0)}{d^{[i]}_j + d^{[R]}_j} \]  

Figure 5. Line segments extracted from two stereo images.

**Table 1.** The results of the (a) pair-wise matching and (b) inferred line matches from the pair relations.

<table>
<thead>
<tr>
<th>Pair-wise Matches</th>
<th>Line Matches</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Left</strong></td>
<td><strong>Right</strong></td>
</tr>
<tr>
<td>1–2</td>
<td>a–b</td>
</tr>
<tr>
<td>1–3</td>
<td>a–b</td>
</tr>
<tr>
<td>1–4</td>
<td>a–c</td>
</tr>
<tr>
<td>1–5</td>
<td>e–d</td>
</tr>
<tr>
<td>1–6</td>
<td>a–f</td>
</tr>
<tr>
<td><strong>Left</strong></td>
<td><strong>Right</strong></td>
</tr>
<tr>
<td>1</td>
<td>a–e</td>
</tr>
<tr>
<td>2–3</td>
<td>b</td>
</tr>
<tr>
<td>4</td>
<td>c</td>
</tr>
<tr>
<td>5</td>
<td>d</td>
</tr>
<tr>
<td>6</td>
<td>f</td>
</tr>
</tbody>
</table>

Figure 6. A problematic case of the Redundancy measure.

have two candidates for matching (two different red lines in right column). The blue lines demonstrate the lines that assist the pair relations for both candidates. As expected, the correct match (Fig. 6a–b) had successfully paired with a total of four lines that belong to the surrounding boundaries of the building roof: However, surprisingly, the wrong candidate (Fig. 6c–d) paired with a total of six lines (some of them are multiple matches) extracted from the boundaries of a car parked on the nearby street. Thus, for this example, blindly counting the number of occurrences may lead the redundancy measure to a wrong match (red lines in Fig. 6c–d). Therefore, we weight all pair relations proportional to their within pair minimum distances. By this way, the redundancy measure provides a possibility to understand and integrate a local matching support for lines. It is clear from Fig. 6a–b that the minimum distances between the lines in pair relations that belong to the correct match are much shorter than the ones that belong to the wrong match (Fig. 6c–d). Thus, we propose the new redundancy measure (SimR) for a line pair as:

\[ Sim_R = \frac{1}{N} \sum_{i=1}^{N} \frac{\min(j \mid d^{[i]}_j \neq 0 \text{ and } d^{[R]}_j \neq 0)}{d^{[i]}_j + d^{[R]}_j} \]  

provided that the \( d^{[i]}_j \neq 0 \text{ and } d^{[R]}_j \neq 0 \). In Eq. 6, \( N \) is the number of pair relations assist to matching, \( d_{ij} \) is the pixel-based minimum 2D Euclidean distance between two lines \((l_i \text{ and } l_j)\) in a pair, \( L \) and \( R \) indicates the pair relations in left and right images, respectively.

2.2.3 Pair-wise Quality Measure

During pair-wise matching, the final pair matches are assigned after a weighted pair-wise matching similarity score which is computed over a total of eight measures; an epipolar, three geometric, two photometric, a correlation and a spatiogram measure (Ok et al., 2010). All the measures are normalized from 0 to 1 prior to the calculation, and the total similarity \( (\Theta) \) result is computed as the average of all similarities. The final pair-wise similarity value (between 0 and 1) for each pair that is computed from those eight measures may give us a hint about the quality of the line matches in that pair. Thus, if a line match is a part of \( N \) number of pairs, the pair-wise quality metric \( \text{(Sim}_{Q} \) for that line match is computed as the average of all pair-wise similarities:
\[ \text{Sim}_Q = \frac{1}{N} \sum_{i=1}^{N} \theta_{i}^{[q]} \]  

where \( \theta_{i}^{[q]} \) is total pair-wise similarity of the \( q \)th pair relation.

2.3 Post-Processing

In this paper, we propose a new post-processing algorithm that mainly relies on the pair-wise matches. We integrated the measures explained above for the final disambiguation process in an exclusively developed iterative pair-wise manner, in which the matching inconsistencies are eliminated using nearest/next distance ratios (NNDR) and a final similarity voting scheme.

Our aim during post-processing is to first eliminate indisputably wrong relations based on a very strict NNDR values forced jointly over the measures Daisy (\( \text{Sim}_D \)) and Redundancy (\( \text{Sim}_R \)). NNDR is first introduced by Lowe (2004) based on the fact that the correct matches need to have the closest matching similarity significantly closer than the closest incorrect match to achieve reliable matching. For false matches, there will likely be a number of other false matches within comparable matching similarities. For a large number of datasets, we investigated the NNDR metric in terms of the ratio of closest to second-closest matches of each line, and for line matching problem, we reject all the related matches of a match that has a Daisy dissimilarity (1- \( \text{Sim}_D \)) ratio lower than \( \text{rattio}_D \). The threshold is selected in a way that only a very limited number of line matches that have enough confidence has possibility to fulfill this threshold (\( \text{rattio}_D = 0.1 \)). Moreover, on the contrary to the other studies that rely on a single measure, we also jointly force the redundancy NNDR metric during this process, thus, we also restrict the elimination of matches which has a redundancy distance ratio lower than \( \text{rattio}_R = 0.35 \). Thus, by means of this joint restriction, we eliminate all the indisputably wrong relations beforehand without removing any of the correct matches. Thereafter, we delete the line relations indicated by NNDR from the pair-wise matches, and for each match, we update the redundancy and quality metrics. At this point, it should be pointed out that, if a line relation in a pair is found to be wrong, we do not directly delete the pair, since we don’t have any inference (correct or wrong) for the other match in the pair. The example given in Table 1a clarifies this fact. For the fourth pair relation (1–5, e–d), assume that we found that the line match (1–e) is wrong. However, we do not have any information about the other match (5–d) in the pair; thus, we cannot directly label the other match as wrong (although it may be in some cases). Therefore, since one of the matches in a pair is labelled as wrong, we update the redundancy (\( \text{Sim}_R \)) and quality measures (\( \text{Sim}_Q \)) of the other match in that pair by eliminating the contribution of that pair from its similarity values. By this way, for example, the match (5–d) given in the Table 1a is not directly eliminated, but penalized, due to reason that the relation (1–e) in the pair is labelled as wrong.

Once all the measures are updated, we initiate an iterative matching scheme by starting from the match that has the highest redundancy measure. Subsequently, we select all the potential matching candidates (ambiguities) for that match. Thereafter, for those matches, we compute an overall similarity metric by taking the weighted linear combination of the similarity measures:

\[ \text{Sim}_F = w_D \cdot \text{Sim}_D + w_Q \cdot \frac{\text{Sim}_Q}{\text{Sim}_R} + w_R \cdot \text{Sim}_R \]  

In Eq. 8, for each selected match, we normalize the Redundancy measure (between 0 and 1) with the maximum Redundancy value of the selected matches, so that the contribution of all similarities is consistent for the final voting. Based on our experiments, we found that the redundancy is the most reliable and unique measure among the three measures, thus, in this study, weights of the similarities in Eq. 8 are designed as \( \{w_D, w_Q, w_R\} = \{1/4, 1/2, 1/4\} \).

Apparently, among the selected matches, the correct match is the one that maximizes the overall similarity metric (\( \text{Sim}_F \)). Thereafter, we fix the correct match and check for the matching ambiguities that violate the selected match. At this point, the colinearity of the line segments of the matching violations (if there any) are individually tested with the line segments of the correct match in order to avoid the deletion of the fragmented lines. The ones that are found to be collinear are labelled along with the correct match for the final matching list. The ones that are not collinear are deleted from the pair-wise matches. After the deletion, we apply the same updating strategy as we explained above. Thus, at the end of each iteration, we penalize all related matches in the pairs that are labelled as wrong. Thus, the (updated) measures (\( \text{Sim}_Q \) and \( \text{Sim}_D \)) turn out to be more and more reliable after each iteration.

Finally, the iterations stop after there is no ambiguity exists in the final matching list. Like any other system developed so far, when a line segment in the first image has no corresponding line segment in the second image, the system cannot identify the wrong match (if accidentally assigned) since the correct line to be matched is missing. To solve this problem, a final check with a global threshold is required. On the contrary to the previous studies that rely on a single threshold, we propose a new hysteresis like global thresholding to solve the problem and to maximize the performance of the matching. As we penalize the Redundancy measure (\( \text{Sim}_R \)) for each match after each iteration, once the iterations has stopped, we have a near-perfect final (\( \text{Sim}_R \)) values for the final matching list. This gives us a unique way to solve the above mentioned problem, in principle; those ill-posed matches have very low Redundancy values when compared to values of the correct matches. Thus, we define a two-level global thresholding:

(i). \( \text{Sim}_D \geq \text{Thr}_D^1 \)  
(ii). \( \text{Sim}_D \geq \text{Thr}_D^2 \) & \( \text{Sim}_R \geq \text{Thr}_R \)

From experiments, we have found that a global Daisy threshold of \( \text{Thr}_D^2 \geq 0.2 \) must be independently satisfied by every match. However, due to lack of rich textures in line local neighbourhood, some false matches may easily exceed this threshold. Increasing the threshold may have a possibility to eliminate some of the correct matches as well, thus results in reducing the overall completeness of the matching. So, we propose to utilize a second high Daisy threshold, \( \text{Thr}_D \geq 0.85 \) restricted with a global Redundancy threshold of \( \text{Thr}_R \geq 0.5 \). By this way, compared to case where only a single global threshold is forced, using a two-level thresholding at the same time can eliminate most of the remaining false matches while keeping the matching precision and matching completeness.

3. RESULTS AND DISCUSSION

To test our methodology, we selected four urban test sites with various built-up characteristics (dense-sparse, flat-gable-complex roofs, repetitive patterns etc.) over a built up area of the city of Vaihingen–Germany. The stereo pairs were acquired by the DMC digital camera with 70% forward overlap (Cramer
The focal length of the camera was 120 mm and the flying height was approximately 800 m above the ground level, which corresponds to a final ground sampling distance (GSD) of approximately 8 cm.

For all test sites, we applied a 50 m (= 162 pixels) search range difference (between the min. and max. heights) along the epipolar lines. The number of correct and false line matches was assessed manually, and first and second columns of Fig. 7 show the matched lines for the left and right stereo images, respectively. For all test sites, more than 55% of the extracted lines are matched, and of these matches 98% are correct. If the complexities of the test sites are taken into account, this seems to be a very good performance. Furthermore, on the contrary to the most of the previous approaches, we do not impose any external dataset to the matching (third view, DSM etc.), and do not perform any ill-posed constraint, such as epipolar ordering.

It should also be emphasized that the curved segments

<table>
<thead>
<tr>
<th># of Lines Extracted</th>
<th>Left</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1823</td>
<td>1877</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1106</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Correct</td>
<td>1088 (98.4%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>False</td>
<td>18 (1.6%)</td>
</tr>
</tbody>
</table>

Figure 7. The results of the proposed approach. Matched line segments are shown in green color in the left stereo images (a-c-e-g) and the right stereo images (b-d-f-h) for the selected four test sites.
(especially the ones belong to the road segments) that can be piece-wise linear approximated are also matched successfully. Actually, this is not a surprising fact, since the piece-wise approximated linear segments are also particularly suitable to be matched by the proposed pair-wise approach.

4. CONCLUSIONS

In this study, we developed a new pair-wise relation based approach for the matching of line features from stereo aerial images. To solve the matching inconsistencies, we proposed an iterative pair based post-processing algorithm. The novelty of this study originates from the newly defined measures and the iterative pair-wise elimination in which the nearest/next ratios and a final similarity voting scheme are applied.

Based on the results of the selected test sites, the proposed approach produces accurate and robust results for urban areas even under challenging cases, such as repetitive linear patterns. We would like to stress once more that, we do not impose any external dataset to the matching (third view, DSM etc.), and do not perform any ill-posed constraint, such as epipolar ordering to solve the matching ambiguities. Thus, our first aim for the future work is to adapt the algorithm into a multi-stereo approach where the third image is fully integrated. For sure, the addition of the third image in a multi-stereo approach will boost the performance of our algorithm in all aspects; accuracy, robustness and completeness. Currently, we are also investigating new approaches that take into account the pair relations for the reconstruction of line features from stereo images. The pair-wise approach also provides new opportunities for solving the most problematic case of the reconstruction in which the matched line segments are exactly aligned with the epipolar line (flight direction).

ACKNOWLEDGEMENTS

The Vaihingen dataset was provided by the German Association for Photogrammetry and Remote Sensing (DGPF): http://www.ifp.uni-stuttgart.de/dgpf/DKEP-Allg.html

REFERENCES


