FUSION OF DATA FROM FUZZY INTEGRAL-BASED ACTIVE AND PASSIVE COLOUR STEREO VISION SYSTEMS FOR CORRESPONDENCE IDENTIFICATION

Alois Knoll, Ralf Schröder, and André Wolfram
University of Bielefeld, Faculty of Technology,
Department of Computer Science,
Postfach 10 01 31, D-33501 Bielefeld, Germany
e-mail:{knoll,andre}@techfak.uni-bielefeld.de

ABSTRACT

As shown in our previous work, an approach using the fuzzy-integral [3] can be applied to solving the correspondence problem of active colour stereo vision systems [2]. Evaluating the similarity measure derived in [2] enables the identification of a correct match or otherwise indicates at least several possible matches. To reduce the remaining ambiguity further, the novel approach presented here uses data fusion techniques to make use of additional fuzzy feature-based information gathered by passive colour stereo procedures. Our experimental results, which are discussed in the paper, indicate that this new approach is considerably more effective than the approach using only intensity-based information for determining the similarity of line blocks in colour stereo images.
We conclude the paper with a discussion of the potential of the method and directions of possible future research.

1 INTRODUCTION

In general, passive stereo algorithms are more reliable in matching pixels in the vicinity of features such as edges than matching pixels belonging to large and unstructured regions. By contrast, active procedures as developed in [2] use colour codes projected onto the objects in the scene to identify corresponding pixels even in featureless areas, but suffer from deficiencies pertaining to the identification of correspondences in the neighbourhood of sharp edges that produce occluding contours. It is therefore desirable to combine the advantages of both approaches. Since the active system requires much less computational effort it is used as long as corresponding colour codes can be found in the scene. Once the measurements become erratic or ambiguous, the algorithms for the evaluation of the data of the passive system are activated and combined with the uncertain active results.

Our experiments show that whenever an ambiguous match occurs, the supporting passive stereo procedure is frequently capable of resolving the ambiguity, in particular if the regarded pixels are located in the vicinity of characteristic features. In this way fusion between data of active and passive stereo algorithms can result in a more dense range map than could be achieved by the single procedures.

2 FUSION STRATEGIES

In order to show the benefit of the fusion of data acquired from active and passive procedures, a method is presented which combines range data computed from active stereo based on colour hues and feature-based passive colour stereo. In the former, a spatially changing colour pattern is projected onto the object and pixel-correspondences are found by comparing colour hues, while in the latter "straight edges" along with some attributes of these edge lines are extracted from the stereo images and are used as features for the correspondence search.

Both the active and the passive stereo procedures comprise the following processing steps:

- **Acquisition of image data.**
  Two colour cameras are used for recording the images of illuminated objects (in the active case) or objects reflecting only the ambient light (in the passive case).
- **Low-level processing.**
  Determination of the colour code for any regarded pixel (active case); estimation of gradients in the neighbourhood of each pixel and determination of the angular deviation among these gradients. Based on this information identification of points in the vicinity of edges is performed (passive case).
- **Extraction of line segments.**
  Computation of epipolar lines that contain the corresponding pixels (active case); extraction of straight edge segments composed of the edge points mapped in the previous stage (passive case).
- **Correspondence search.**
  In the active procedure the colour codes of pixels along corresponding epipolar lines are compared, while the passive procedure compares the attributes of extracted straight edges.

- **Generation of range maps.**
  For related pixels, i.e. those marked as corresponding in the previous step range values may be calculated and visualized in a range map.

Data integration of active and passive stereo algorithms may take place on various levels for different purposes:

- **Filtering the input data of the active stereo procedure.**
  For the purpose of generating range values for contour edges of an object, the input data of the active procedure may be restricted, for example, to those pixels labelled by the passive procedure as being located in the vicinity of an edge (the active procedure supports the passive procedure in order to examine boundaries more closely). Alternatively, the active procedure is applied only to those regions that are limited by continuous lines (regions are examined more closely).

- **Simplifying the search for correspondences.**
  As the illumination of objects is completely different in both cases, the information sets gathered from both procedures are in some sense "orthogonal", i.e. may complement each other. The feature-based passive procedure may profit from additional colour information at object lines, whereas the intensity-based active procedure may benefit from information on edge locations during the colour matching process. Decisions made while the passive algorithm progresses (e.g. image point is close to an edge, for which possibly another correspondence candidate has already been found) can be utilized in an iterative process to resolve ambiguities of the active procedure. Conversely, if corresponding pixels were found by the active procedure lying on (straight) edges, these edges are likely to be corresponding edges.

- **Improving the credibility of the input data.**
  Since there are multiple images of the same scene available and the effects of image noise are different for each image, the signal-to-noise ratio may be improved. Furthermore, edges created by shadows or edges of regions with a different hue may or may not be visible under various lighting conditions.

- **Improvement of the resulting output data (correspondence map).**
  Data gaps resulting from one procedure may be eliminated from the map making use of the results produced by the other method. It is obviously desirable to mutually check the results of the output of both procedures and to register it in the correspondence map only if the distance between the results remains below certain threshold.

We now concentrate on our main goal, i.e. the generation of dense range maps. Fusion methods aiming at improving the credibility of input data or at restricting the working area of the active procedure are not considered further. In the sequel we describe a fusion algorithm that utilizes the same images of colour-illuminated scenes for the active and the passive procedure. Output of the passive method is used whenever the results of the active technique cannot generate unambiguous results.

3 FUSION ALGORITHM

The algorithm searches for pixels in the left and right image with the highest degree of similarity in colour. The search is conducted on corresponding epipolar lines using fuzzy integral-based similarity measures as given in [2]. To determine the similarity Ξ^* between two blocks H_F^{l} and H_F^{r} of adjacent colour pixels centered around the pixels to be compared, fuzzy integration of a suitable intensity-based similarity measure Ξ (such as the commonly used mean square error) over $F' = \{(H_R^{\mathrm{l}}, H_R^{\mathrm{r}}), (H_G^{\mathrm{l}}, H_G^{\mathrm{r}}), (H_B^{\mathrm{l}}, H_B^{\mathrm{r}})\}$ is carried out with respect to a g_λ-fuzzy measure, which models the credibility of the individual colour channels (l/r denoting the left resp. the right image, H_F/H_f a colour resp. an intensity block, $f \in F = \{R, G, B\}$ and $f' \in F'$):

$$\begin{aligned} \Xi^*(H_F^{\mathrm{l}}, H_F^{\mathrm{r}}) &= \fint_{F'} \Xi(f') \circ g_\lambda(\cdot) \qquad (1) \\ &= \sup_{A \in \mathcal{P}(F')} \min[\min_{f' \in A}(\Xi(f')), g_\lambda(A)] \end{aligned}$$

As long as unambiguous matches are found, they are marked accordingly and used to calculate range values. Ambiguous matches may occur during the search for corresponding pixels due to various reasons:

- The fuzzy integral can propagate identical similarity measure values for slightly different colour codes of the compared line blocks. This effect may be reduced, if definitions other than the classical form of the fuzzy integral as given in (1) are used. Examples are the use of other t-norms, such as the algebraic product instead of the minimum-operator.

- Significant changes of depth in the stereo scene may cause several appearances of identical colour codes in the regarded epipolar lines, even if the coloured light pattern itself is non-repetitive.

- The colour of objects in the scene may change the projected light pattern by absorption in a way that several identical colour codes are generated, etc.

In case of such *ambiguous* matches the following steps are taken:

For the pixel in the left picture a list of possible matches along a corresponding epipolar line in the right picture is generated. The colour pattern projected onto the scene ensures that these matches are *local maxima of the pixel colour similarity measure*. Due to the hue of the colour pattern changing continuously, multiple appearences of similar or identical colour codes in an epipolar line are in general separated from each other by less similar codes.

Assuming that the reference point in the left picture and every list entry representing a point in the right picture are actual corresponding points, the range of the object point related to the pair of image points is calculated by triangulation. Modelling a *measure for smoothness* from this data using fuzzy sets seems to be useful because there are no certain criteria for up to which limit a difference in range should be regarded as smooth. Moreover, the average range may vastly differ from the true range at a specific point and therefore could indicate a possible match creating a smooth surface (although it does not). The average range value at the position of the considered reference point in the range map is determined using entries of the range map calculated previously within a certain environment.

Furthermore, gradient information of each pixel is examined. Therefore, all following steps concerning gradient data are processed three times, once for each of the separate colour channels (i.e. $\forall f \in F$):

A measure is calculated for the *angular deviation* of the gradients in the neighbourhood of the reference pixel as well as all possible matches. This measure has been derived from a method being presented in [1]. For each pixel in the neighbourhood of the point of interest gradient vectors are estimated using Sobel operator masks. Summation of these gradients $\mathcal{G}$ and their absolute values $|\mathcal{G}|$ leads to the angular deviation δ indicating the presence of an edge in the vicinity of the considered point if the angular deviation is low with respect to the gradients evaluated (for reasons of simplicity the colour channel index f is omitted):

$$\delta(\mathcal{G}) = 1 - \frac{|\sum \mathcal{G}|}{\sum |\mathcal{G}|} \qquad (2)$$

Regarding the angular deviation for both pixels of a possible match three slightly different cases may occur, leading to the following conclusions:

- **Vastly different values for $\delta(\mathcal{G}^l)$ and $\delta(\mathcal{G}^r)$.** With one of the related pixels being located in the vicinity of an edge while the other is not, the pixels are unlikely to be corresponding.

- **Rather high values for $\delta(\mathcal{G}^l)$ and $\delta(\mathcal{G}^r)$.** Both pixels are likely to be located in image regions without features suitable for verification of possible correspondences. There are no edges to compare and the directions of the gradients vary randomly.

- **Very low values for $\delta(\mathcal{G}^l)$ and $\delta(\mathcal{G}^r)$.** Both pixels are likely to be located in the vicinity of an edge, and thus are probably actual corresponding pixels.

A measure α, which indicates whether a possible match belongs to a pair of corresponding edge points or at least edge region points, may therefore be defined as:

$$\alpha\left(\delta(\mathcal{G}^l), \delta(\mathcal{G}^r)\right) = 1 - \max\left[\delta(\mathcal{G}^l), \delta(\mathcal{G}^r)\right] \qquad (3)$$

Gradients in the neighbourhood of the reference pixel in the left image and of the possibly corresponding pixel are compared to determine a further similarity measure, which is constructed from intersections of the fuzzified components of all gradient vectors concerned. This is done separately for the X- and Y-components, to consider the direction of the gradients as well as their magnitude. Particularly in case the angular deviation indicates the presence of edges in the vicinity of the possibly corresponding pixels, the *measure for the similarity of gradients* can be interpreted as a measure for the similarity of the edges close to the regarded pixels. If, for example, the pixels of a possible match can be identified as being close to edges with significantly differing directions, these pixels are not likely to represent an identical object point.

Subsequently all possible matches are rated by the minima ϵ_{f_i}, $i = 1, \ldots, m$ (with m denoting the number of local maxima of Ξ^*) of the proposed criteria:

- the measure of similarity Ξ^* between colour codes of line blocks, as determined by the active stereo procedure using the fuzzy integral,

- a measure for range map smoothness deduced by intersecting fuzzy sets for the range calculated for a single match and the average range of the map in the neighbourhood of the regarded point,

- the measure α, emphasizing possibly corresponding pixels that are both located on or near an edge and

- a measure for the average similarity of gradients in the neighbourhood of the possibly matching pixels in both images

separately for each colour channel f.

The highest ratings $\hat{\epsilon}_f = \max_i \epsilon_{f_i}$ indicate the possible matches which are most likely to be the actual correspondence considering gradient information of just a single colour channel f. For the final decision we select the match with the overall maximum $\epsilon_F = \max_f \hat{\epsilon}_f$ because a channel, detecting only image noise instead of an edge, will produce a very low rating, or some edges may only be visible in a single colour channel due to different absoption of the light pattern projected onto the scene.

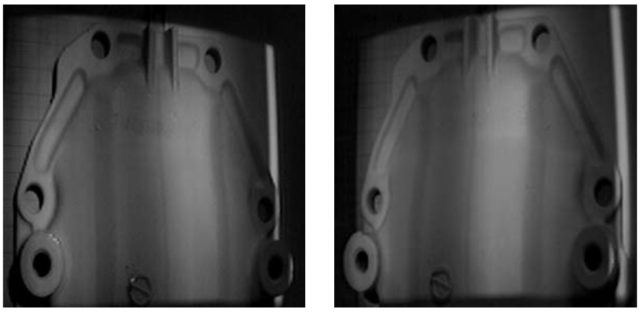

Figure 1: Test scene

That match is then regarded as the new reference pixel in the right picture. The whole procedure of finding the match with the greatest similarity is repeated, comparing this new reference with possible matches on an epipolar line in the left picture. If the pixel with the highest rating, determined after completion of the second run, coincides with the original reference pixel (within a small tolerance), these pixels are most likely to be actually corresponding, and therefore this match is marked as *unambiguous*. Otherwise the status of the regarded pixel remains *ambiguous*.

4 EXPERIMENTAL RESULTS

Evaluating stereo images of a size of 290×272 pixels (a white automotive part, see figure 1), the results shown in table 1 were obtained. The single colour-coded active stereo procedure produces 20258 ambiguous matches; the additional use of fusion techniques generates 19632 more unambiguous matches.

procedure:	active stereo	fusion applied
allocated matches		
pixels overall:	78880	78880
no match:	14536	14536
ambiguous:	20258	626
unambiguous:	44086	63718

Table 1: Experimental results

In order to give a better impression of the spatial relations between the occurences of ambiguous and unambiguous matches the same facts are shown in figure 2. Unambiguous matches appear as white, ambiguous matches as light grey, those produced as unambiguous by fusion are marked as dark grey. Black areas contain pixels for which no matches were found. Figure 3 shows the corresponding range map (brighter grey levels represent shorter distances between the object and the stereo camera system).

5 CONCLUSIONS

The recent results presented in this paper indicate a definite improvement over the pure active stereo procedure by reducing ambiguities in the correspondence search. Fusion on a higher level using the results of passive stereo matching algorithms in addition to the active colour stereo data is an objective of further research. Attempts are also made to find efficient methods to auto-adjust parameters concerning the modelling of fuzzy sets used in the active as well as the passive stereo procedure.

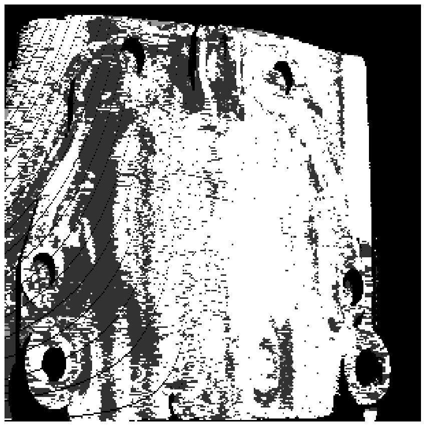

Figure 2: Correspondence map

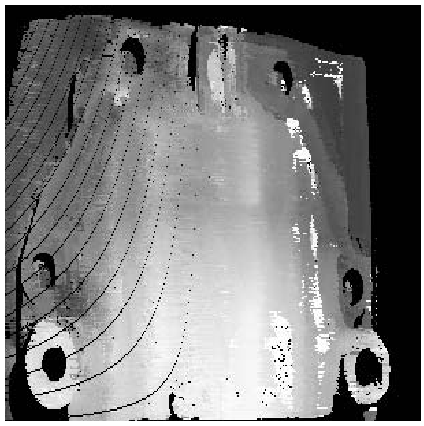

Figure 3: Range map

References

[1] Gregson, P.H. *Using Angular Dispersion of Gradient Direction for Detecting Edge Ribbons.* IEEE Trans. on PAMI, Vol. 15, No. 7, 1993, 682–696

[2] Knoll, A. and Wolfram, A. *Searching Correspondences in Colour Stereo Images - An Approach Using the Fuzzy Integral.* Proc. of the Int. J. Conf. of the 4th IEEE Int. Conf. on Fuzzy Systems and the 2nd Int. Fuzzy Engineering Symp., Yokohama, Japan, 1995, 2033–2038

[3] Sugeno, M. *Theory of fuzzy integrals and its applications.* PhD thesis, Tokyo Institute of Technology, 1974