Matching using object models generated from photometric stereo images

G. Bellaire*, K. Schlüns, K. Oppermann, W. Schimke

Technical University Berlin, Department of Computer Science, Vision Group, Sekr. FR 3-11
Franklinstr. 28-29, 10587 Berlin, GERMANY
{bellaire, karsten}@cs.tu-berlin.de

ABSTRACT

In this paper we present a complete 3-D object recognition system. The task of the system is the separation of different plastic salvage for recycling of goods' packages. In Germany separation is done by hand. This hand separation is expensive and gains bad results, caused by the work circumstances. The system holds the following characteristics: 1) Viewer centered object representation - the used extended surface normal image (ESNI) model is a new and promising representative of this approach. 2) Photometric Stereo Method (PSM) - we have integrated a specular high light elimination and we have developed a method, that combines dense gradient information and a line drawing of the object. 3) An adaptive hierarchical indexing of geometric features is used as a first matching step. Thereby an online adaptation by integrating the sensor accuracy is performed. 4) A main feature of the system is the 3-D rotation invariance of the unique base face orientation achieved by PSM. A color indexing and an intelligent template matching method are introduced to receive the recognition fine-tuning. Both methods are powerful but assume a fixed 3-D orientation of the objects. Because of the achieved (well-defined) base face orientation this restriction is overcome. We show results of experiments with real, deformed and dirty goods' packages.

Keywords: Object recognition, photometric stereo method, extended surface normal image representation, color indexing, adaptive template matching

1 INTRODUCTION

Viewer-centered approaches [1, 2, 3, 4, 6, 8, 9] develop models that represent the information visible from a certain view point. The projective relation between the 3-D object and the image is not considered during the matching process. The presented system uses the extended surface normal image representation (ESNI), which is derived from the surface normal image (SNI) representation [9]. We introduce an adaptive hybrid matching framework and show results with real data in the

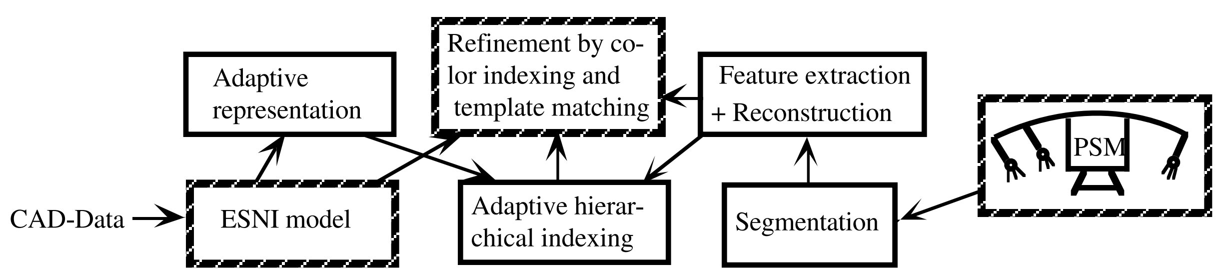

Figure 1: Instances of our object recognition system. The highlighted components are presented in this paper.

context of waste package separation. The task of the system will be the separation of different plastic salvage for recycling of goods' packages.

In Germany the goods producing industry is obliged to recycle packages. Recyclable packages carry a logo the so-called green point. These packages are sampled in particular recycling bins, that are distributed in the private households. Specialized companies have to extract a certain percentage of recyclable packages, which is prescripted by law. In Germany separation is done by hand. This hand separation is expensive and gains bad results, caused by the work circumstances. Our method combines object- and data-driven strategies to a hybrid system. The data driven indexing generates hypotheses that are localized and verified by an object-driven color indexing and an analysis of the package imprint. Our work focuses the following aspects:

1) Computing the ESNI representation from CAD-data and from the scanned package imprints: We use a self developed Mini-CAD system as an interface to the well-known CAD-format DXF. The ESNI integrates geometric data, data suitable for our color module and the image information of the processed objects.
2) Developing an adaptive index layer model: The exact classification of scene objects and the efficient access to complex model data - largely independent of the number of objects- are goals of a recognition system. The use of "a priori" knowledge about the feature distribution in the data base supports the recognition realization. Further, an interface to an "a priori" analysis module of the actual sensor configuration is put at the user's disposal [3].
3) Improving the indexing by integrating the measured sensor errors: We developed the adaptive hierarchical indexing and its learning module, that adjusts the indexing in dependence on the sensor behavior.
4) Fine-tuning by using color and image texture: To process waste packages we use a color indexing process, that overcomes its principal restrictions by using results of the Photometric Stereo Method (PSM). Additionally we use a ‘3-D template matching‘ to improve the results of the color indexing.
5) Presenting an albedo independent shading based method for reconstructing surface gradient images which is able to process colored objects.
6) Reconstructing a 2.5-D model from surface gradient images.

Aspects 1, 4 and 5 are described in this paper. The aspects 2, 3 and 6 are described in [2].

2 RELATED WORK

Park [9] originally introduced the SNI representation. He presented a recognition system which uses range images as image data. Park did analyze geometric features without concerning texture or color information. Ikeuchi [6] presented a system for the automatic generation of recognition and localization programs. He investigated the linear view changes in characteristic view classes by employing decreased sets of characteristic views, which correspond to certain sensor configurations. Ikeuchi although did only deal with geometric features. Flynn [4] described the transformation of CAD-object representations to viewer position independent relational graphs, that are appropriate for vision applications. The uniform tessellation method generates 2-D views. Lu [8] applied CAD-data to generate a vision model and a recognition strategy. A data structure called relational pyramid manages primitives of different complexity levels. To process inexact matching they have investigated the accumulator based matching, that adds votes to views with similar feature constellations. Swain and Ballard [13] originally introduced the color indexing approach to match and localize 2-D objects in multi object scenes. The Photometric Stereo Method was introduced by Woodham [14].

3 3D SHAPE RECOVERY USING PHOTOMETRIC STEREO

Shading based methods for surface reconstruction consider the imaging sensor as a photometric measuring device. Assuming special reflection characteristics and known illumination configurations, the measured image irradiances (intensities) reduce possible surface orientations to a closed curve on the Gaussian sphere. For Lambertian reflection these curves are circles. Several appproaches exists to achieve a unique surface in spite of this ambiguity in image irradiance.

Most approaches require hard limitations to the object world. This implies that these approaches are theoretically very interesting but not useful in practice. However shading based methods have the important advantage that the three-dimensional features they calculate can be recovered in a dense manner. Therefore, a surface orientation or a relative depth value can be assigned to each visible surface point in the image, except for sharp discontinuities.

In our work we focus on methods which are independent from the actual surface albedo. The reason are practical considerations that come from our application. Albedo independence allows the surface to be of an arbitrarily colored texture. Hence, no reflection factor measurements are needed for the original unknown surface.

Using at least three light sources, surface orientations together with albedo values can be measured indenpendently for each image point (Photometric Stereo Method [5, 14], PSM). PSM calculates surface shape (orientations) from the shading variations in three images, taken of the objects with the light source in different positions. In opposition to Binocular Stereo techniques, the visual sensor is fixed in the same position and orientation. Therefore no matching or registration strategy is necessary.

From the surface orientations the depth can be calculated using an integration technique. Comparisons with several approaches have shown, that a FFT-based method [5] produces the best results.

Applying three-source PSM to a large number of different objects, it has turned out, that this shading based approach can serve as a good starting-point for high-quality surface reconstruction. The investigations have shown, that three-source PSM is open to improvement in several ways. For the application in waste processing the three most important items are:

- using more general reflection models,
- independence on inaccurate illumination parameter estimation,
- and reconstruction of polyhedral objects.

All items concern the accuracy of the surface reconstruction. Some items concern furthermore practical considerations, i.e. ways have to be found how the applicability of the calibration and reconstruction phase can be improved.

3.1 Highlight elimination

Traditional three-source PSM restricts the reflection characteristics of the analyzed surface to Lambertian reflection. This holds for many materials. But in general, surface reflection is composed of a diffuse and a specular component. With the specular reflection component more parameters must be known in advance to model this hybrid reflection. These are roughness values and weighting factors, for example. Using color is a good way to get more information about the scene without introducing more light sources and without making stringent restrictions to the surface reflection properties. We make use of the Dichromatic Reflection Model [12, DRM]. The DRM can be employed to separate the reflection components from a single RGB-image [7, 11]. Moreover, roughness values and weighting factors can be extracted from the images. Since PSM additionally recovers surface albedo values, illumination independent surface color descriptors can be obtained. This is important for the matching process described in Sect. 7.

The methods developed in [7, 11] assume that the number of different materials (body reflection colors) in the scene is relatively small. For example, such approaches fail if object surfaces possess color transitions. Using such methods in waste processing this would be a strong restriction. If separation of reflection components is applied to a three-source PSM environment, then we can take advantage of the fact that three RGB-vectors are measured for each image location. Image locations that can not be handled with the above methods are processed by a local method. The local method assumes that at least one of the three RGB-vectors contains no specular highlight component. Such a RGB-vector can be found by calculating the angle between the known illumination color and the respective RGB-vector. Then the RGB-vector with the largest angle is expected as containing only a diffuse reflection component. This vector can be used for projecting the other two color vectors onto it. The projected colors have no longer a specular reflection component.

3.2 Photometric Stereo analysis with table look-up

In traditional three-source PSM the accuracy of orientation determination depends strongly on the estimation of the light source parameters, i.e. the light source strength and the illumination direction. On the other hand Look-Up Table (LUT) techniques exist that need no knowledge of these parameters, since the LUT is built with a calibration object [5]. LUT techniques in general have the disadvantage that only objects with surface materials can be analyzed which have exactly the same reflection characteristics as the calibration object. This implies many practical problems, since 1. often it is difficult to construct a calibration object with a special surface material and 2. the albedo of the analyzed object must be constant. We have developed a LUT technique which is independent on the surface albedo. Nevertheless the light source parameters leave unknown. In comparison with other LUT techniques the further assumption is, this the sensor has a linear characteristic. The technique is independent of the camera iris setting, since surface orientation computation for a local image position does not depend on the irradiance-vector built by the three measured irradiances. That is important because CCD-cameras have a limited dynamic range and therefore sometimes irradiances are too low for processing if a global constant iris setting is used. For each color channel an own LUT is computed since the spectral composition of the colors often varies strongly.

3.3 Reconstruction of polyhedral objects

For generating polyhedral reconstructions the first partial derivatives generated from PSM have to be transformed to a height map or a geometrical 2.5-D model. We have developed a method [2], that combines dense gradient information and a line drawing of the object. The method consists of two steps. The projected object is taken apart in planar and curved patches. Then these parts are fitted together to a 2.5-D surface. First the gradient images are segmented with region growing. Using this kind of segmentation technique planar as well as curved patches can be extracted without additional effort. Curved patches are approximated by planar patches. These planar patches are attributed as belonging to a curved region. This is necessary for the treatment of occlusions and the elimination of approximation edges in the recognition part. Since region boundaries in the interior of curved patches are determined by the growing process, these patches are post-processed with a balancing algorithm.

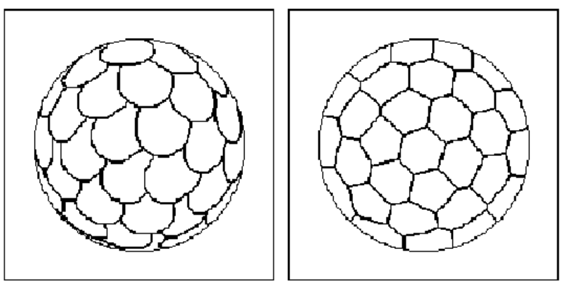

Figure 2: The segmentation of a sphere is exemplary demonstrated. On the left-hand side results of the pure region growing are shown. On the right-hand side results of the balancing algorithm are shown.

A balancing algorithm fine-tunes the extracted regions. Therefore the surface normals of the boundary points of the extracted regions and the mean surface normal direction are compared. A considered boundary point is reclassified if the average orientation in an adjacent region has a smaller angular deviation than the original region. The algorithm is recursively applied until an equilibrium is reached. An equilibrium is determined if no boundary points are assigned to their neighbored regions. In Fig. 2 the left image presents the result of the region growing and the right image shows the result of the described balancing algorithm. Consequently, for the next steps consistent region boundary information is available. A winged edge model is generated from the region and boundary data, details see [2].

4 MODELING BY EXTENDED SURFACE NORMAL IMAGES (ESNI)

The ESNI model is a new and promising approach that uses viewer-centered representations to model 3-D objects. The ESNI model contains as the characteristic view model [3, 4, 8] 2-D views of the represented object. In comparison for most objects the SNI representation contains fewer views as the ChV representation. The SNI model assumes orthographic projection. Additionally the ESNI representation includes color and texture information.

The geometric part of each view is represented as a 2-D line drawing. This geometric part of the ESNI consists of a SNI set. In the SNI set each view is a normal view, which is defined by the alignment of the surface normal of a base face parallel to the line of sight. A SNI model of a 3-D object contains each object face as base face exactly once. In the case of coplanar faces the same SNI is generated several times. These SNIs are not deleted, since the determination of the coplanarity attribute in real gradient images is not a trivial task. It is also possible that a base face is partially occluded caused by self shadowing. The SNI generation algorithm proceeds as the following: Surface normals of the base faces are converted into view points on the Gaussian sphere. SNIs are generated by taking snapshots of the calculated view points. An algorithm to calculate the hidden lines and hidden surfaces is applied for the snapshot generation. The ESNI representation combines the image information of each face and a color histogram for each face with the SNI structure. Fig 3 shows the image and histogram part of an ESNI representation.

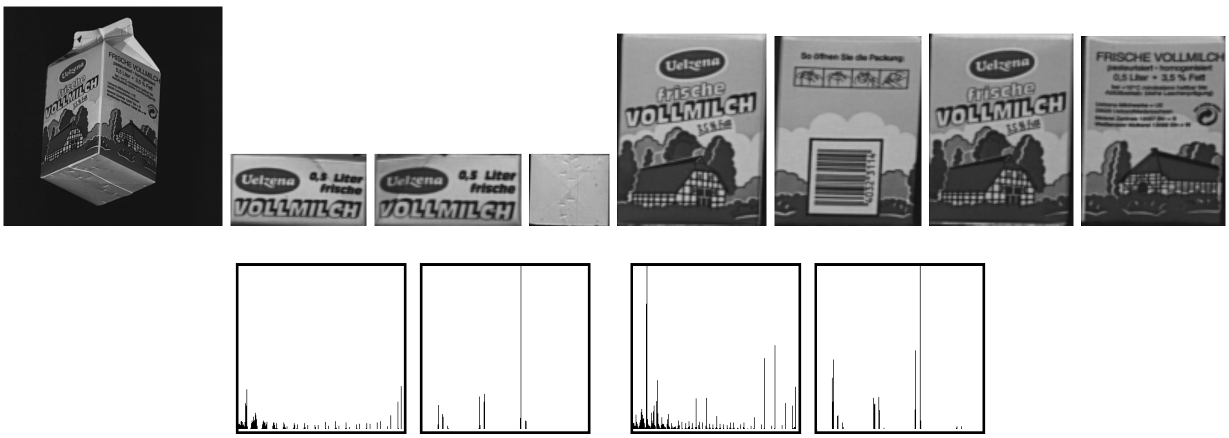

Figure 3: Extension part of the ESNI representation of a milk package. Seven package imprints and the HS and $I_1 I_2 I_3$ histograms of face '2' und '4' are shown.

In the scene domain, the geometric information of an ESNI, which is represented in the SNI part of the ESNI, is matched with so-called rotated input images (RII). The RIIs are generated from a 2.5-D scene model in the same way as the SNIs. Following these descriptions the demand of an especially developed matching strategy is obvious.

Fig. 4 shows a planar object and the representing SNI set. Looking at the SNI number '4' only the base face '4' is visible caused by the 90° angle between the base face and its adjacent faces. The SNI number '2' consists of three faces: the base face '2' and the adjacent faces '1' and '3'.

Fig. 5 presents the SNI set of an object that partially consists of curved faces. In the SNI model curved faces are represented by virtual planar patches. The curved object face '4' of the graphic is represented by six SNIs whereby each planar patch becomes a base face exactly once. The number of the visible patches and their 2-D region size in a SNI depend on the position of the base face patch on the represented surface. In the following we discuss the dependence of the SNI number on the granularity of the virtual tesselation of the curved faces. The curved face of the object of Fig. 5 is represented by six patches, whereby each patch contains surface normals variing over an interval of the size B = 30°.

The number of the SNIs, that represents an object, depends on the number of planar faces and the number of curved faces muliplied by a factor F. F has to be determined for the SNI data base. It consists of: 1) The value of threshold B that determines how many surface points are clued together depending on the variation of the surface normals in these points and 2) The variations of the surface normals in the processed object faces, $V\Theta$ and $V\varphi$, whereby $\Theta = [\min\Theta, \max\Theta]$, and $\varphi = [\min\varphi, \max\varphi]$, with Θ, φ, as sphere coordinates of the surface normals.

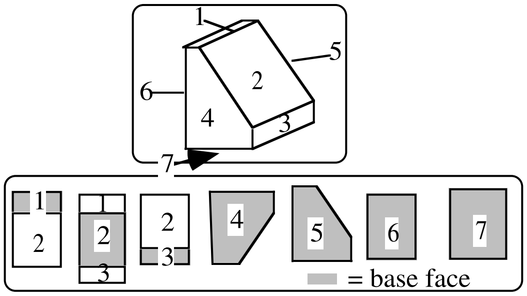

Figure 4: SNI of a polyhedron.

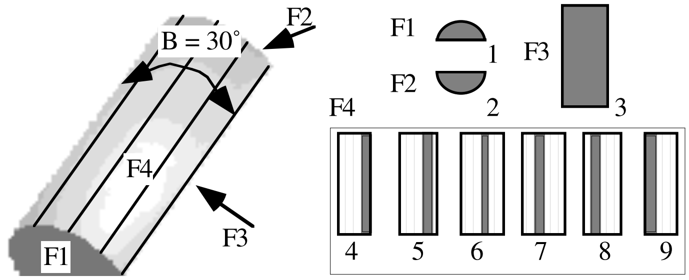

Figure 5: SNI set of a partially curved object.

Following this, F is determined by $F = V\Theta\ /\ B + V\varphi\ \ /\ B$. Fig. 6 exemplary shows the parameters on a curved face. The total amount of the SNIs with PF representing the number of planar faces and CF the number of curved faces, is described by

$$PF + \sum_{i=1}^{CF} F_i\ .$$

During the recognition process the extended rotated input image (ERII) scene representation is generated from the 2.5 reconstruction of the scene object. The ERII is only a subset of the ESNI since we are only able to recover a 2.5-D reconstruction and faces on the object's back cannot be reconstructed.

A RII set is derived from the 2.5-D reconstruction. The approach is similar to the generation of the SNI set. There are some distinctions between the two approaches, since the RII generating algorithm is applied to real 2.5-D data and the data base objects used by the SNI generation are modeled as 3-D CAD-data. The RIIs are a subset of the SNIs that represents the data base object. Fig. 7 shows the SNIs of a polyhedral object and the RIIs generated from our PMS-based reconstruction. The RII set consists maximally of half the number of the data base SNIs.

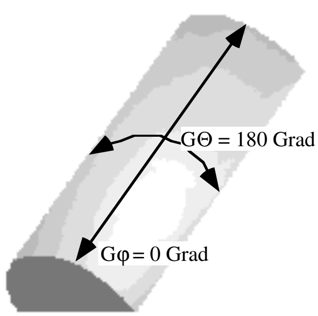

Figure 6: The number of the base faces of a curved surface depends on the threshold B that determines how many surface points are clued together in an approximation patch and on the curvature of the processed surface.

Missing faces in the RIIs are another problem for the matching. In some cases the 3-D rotation of a 2.5- D reconstruction causes visible faces to be turned away from the viewer. Since faces situated on the back side of the object are not reconstructed, some of the RIIs consist of a subset of the faces, that determine the corresponding data base SNI. We are using CAD-models of the processed objects, that are rough approximations of the real world objects. Additionally the used segmentation algorithm produces an over segmentation in case of object details. Following this, only faces with a size larger then 10 percent of the size of the largest reconstructed face are used for the RII generation. It is the main challenge of the implemented matching algorithm to establish correspondences regarding the strong distinctions between the representations.

The 2.5-D reconstruction consist of 3-D coordinates of the object vertices, the reconstructed edges and the object faces. The data format is compatible with the format of the used mini-CAD system. For all faces a representative surface normal is selected by computing the mean of all face normals. The resulting RII depends on the accuracy of the extracted surface normals. Especially in the case of 90 degree angles between the actual base face and the remaining faces small changes of the surface normals cause inaccurate results. In the described case, faces that are not visible in the corresponding SNIs can become visible in the RIIs.

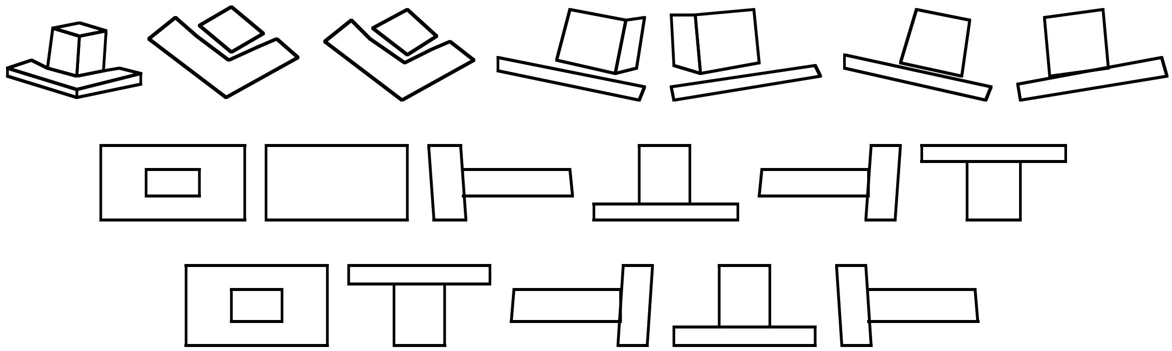

Figure 7: Top row: Six RIIs represent the scene object placed on the left. RII number 3 and 4 contain wrong faces caused by small errors in the extracted surface normals. Bottom row: Corresponding SNI.

5 ACCUMULATOR BASED ADAPTIVE INDEXING

In [2] we introduced the adaptive inexact indexing of geometric features. It gains an optimal matching scheme, that is adapted to the actual object data base, the a priori knowledge about the current sensor configuration, see [1], and the measured sensor errors. The SNI sets of all model objects build the geometric part of an object data base. Feature extracting functions generate feature vectors that describe the particular SNIs. One feature vector exists for each SNI. During the matching procedure the RII set is processed by comparing RIIs' feature vectors one by another with the data base SNI feature vectors. The matching algorithm of the geometric features consists of an adaptive hierarchical indexing scheme, that contains an inexact matching module and that is enlarged by a learning module. Since the indexing has to deal with inexact data, a method has to be found that relates the data to comparative models. An index layer exists for each feature. The proposed matching strategy increases the accumulators (AC) of several SNIs by the value e-k2/2, with k = 0, …, kmax denoting the deviation of a model feature to the measured data. That means the accumulator of a SNI with a feature value close to the measured data is increased more than an accumulator of a SNI with a less similar feature value.
In the proposed system a list of RIIs is compared with the object data base. This process calculates a list of the most voted SNI for every RII. Some regulation has to be done to conclude the particular lists into a single object ranking list. For details see [2]. As shown in [2, 3] the shape of the weighting function and the size of the permitted feature distance between SNI and measured feature value are the most significant parameters to control. In the described system an unsupervised learning component is introduced. The idea of the learning is the usage of gained matching experience to improve the results of the indexing.
The sensor accuracy is measured by the comparison of the correct result feature vector x_{SNI} and the feature vector x_{RII} calculated from the scene object. For each feature, that is element of the feature vectors, the maximal difference between the real and the measured feature value is stored in a learning map. This learning map supervises the mean standard deviation and maximal deviation value max_{dev} of the particular features. After some successfully processed recognition tasks the above mentioned parameters adapt.

6 DIRECT MATCHING OF ALIGNED IMPRINTS

The proposed direct matching compares the imprints of the data base objects with the segmented faces pixel by pixel. The task 'matching of arbitrarily positioned textured faces' has to be performed for the package imprint comparison. The algorithm has to analyze *corresponding* pixels.
The PSM approach calculates the surface normals of the faces of the scene objects. The correct determination of the surface normals is the necessary precondition to perform the direct matching. A pixel by pixel comparison of two faces, which are

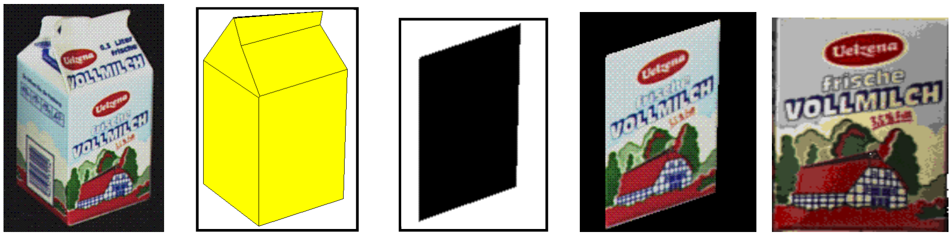

Figure 8: 2.5-D Reconstruction and alignment of the base face shown at the example of a milk package.

arbitrarily positioned, is not possible without an adjustment of the position of the processed faces. Therefore the faces have to be aligned in space. The alignment process is performed by the following steps: At first the face of the data base object is rotated to map its surface normals on the surface normal of the face of the scene object. Then the two longest face edges are aligned on each other. The alignment consists of a rotation around the surface normal and a translation. In the case of similar edge lengths, all possible rotations of the faces are performed and matched. Fig. 8 shows an example procedure of the described steps. A final adaptation of the alignment of the two faces is calculated by the following:
In each data base face an adaptation region with 30 x 30 pixels is determined. This region shows an appropriate amount of contrast. The adaptation region of the data base face and the corresponding scene object face are mapped by an algorithm that calculates the translation vector by finding the position of the most corresponding 10 x 10 window in the region. This translation vector is used for the final alignment of the two faces. Fig. 9 shows exemplary an ideal case of the adaptation. In the figure two equal imprint patches are matched. The patch with the black border represents the adaptation region in the SNI base face. A similar region is placed on the face on that the SNI face has to be aligned. The vector indicates the result of the algorithm. Local displacements are calculated by repeating the process iterativily. Problems of the algorithm are distortions of the reconstructed faces of the scene objects and accuracy changes during the determination of the surface normals.

7 COLOR INDEXING OF WASTE PACKAGES

Swain and Ballard [13] originally introduced the color indexing approach to match and localize 2-D objects in multi object scenes. In our approach the task is to recognize single 3-D objects, that are even partially polluted and deformed. The color information of the package faces is represented by color histograms. Each object face of the data base is represented by one histogram for alle three channels. In this matching step the data base color histograms and the online calculated histograms are compared. The histograms are a part of the ESNIs as shown in Sect. 4. Since we use the same lighting conditions for the data base generation and the scene acquisition, the approach can abstract from color constancy effects.

Figure 9: Final adaptation of the imprint alignment.

For the comparison of the histograms we use the histogram intersection method of Swain and Ballard. In contrary to the original approach the used histograms are relative. In a relative histogram bins are normalized by the total number of pixels The algorithm summarizes the minimal bin value of the two compared histograms. The value of the intersection yields to generate a sorted list of the best matching faces.
We analyzed three different color models to improve the results. RGB and $I_1I_2I_3$ show better results than HSI. In the HSI model the hue value contains to much variation to match the correct bins.
We also analyzed distinct resolutions to discretisize the color spaces. Tab. 1 and 2 show two resolutions of the $I_1I_2I_3$ color space we used. We show results with a resolution of a 4 bit red-green, 4 bit blue-yellow and a 3 bit intensity channel. This resolution gains

2048 bins which need a storage space of 8K per histogram. Additionally we show results of a 3-3-2 resolution which gains in 256 bins and a storage space of 2 K.

Histogram resolution	4-4-3	3-3-2
recognized	72%	66%
among 3 most similar objects	93%	86%
not recognized	7%	14%

Table 1: Polyhedral objects.

4-4-3	3-3-2
71%	43%
86%	86%
14%	14%

Table 2: Dirty and polluted objects.

In this stage of the implementation the algorithm take care about the histograms of the one or two largest faces of the scene and all data base faces.

8 RESULTS

Image aquisition parameters: For image generation we have used a 3CCD-Color-Camera Sony DXC730P. Gain control and gamma correction were turned off. Automatic white balancing was used. The digitization process was done by a DATACELL s2200-Framegrabber in a Sun IPX workstation. The focal length used with PSM was ca. 60mm. The scenes were illuminated by slide projectors. For most objects two iris settings were necessary.

2.5-D Model reconstruction: In the following we show that the LUT technique (Sect. 3.2) works in polyhedral object reconstruction described in [2]. The parameters shown in Tab. 3 are only given to have an impression of the illumination configuration. They are not used to determine surface orientations.

Source No.	Relative Strength	Estimated p_s	Estimated q_s	Estimated Tilt	Estimated Slant
1	1.0	-0.311455	-0.231252	-143.41	21.20
2	1.02206	0.048908	0.303985	80.86	17.11
3	0.77149	0.410855	-0.235918	-29.86	25.35

Table 3: Light source parameters used for generating the 2D model of the milk package.

In Fig. 10 one of the input images and a needle map of the reconstructed surface gradient image are shown. The cow drawn on the package has a very small albedo. Therefore (among other regions) no surface orientations are determined in the cow since in this case only a single iris setting was used.

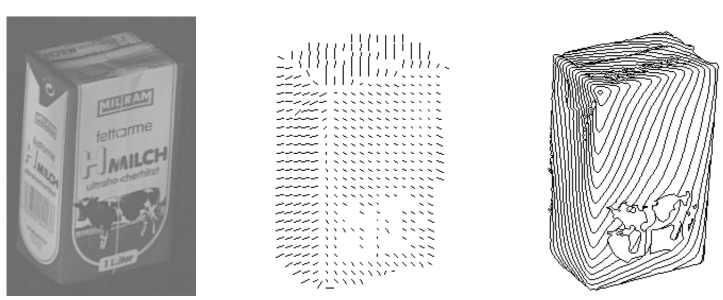

Figure 10: PSM input image of the milk package, a needle map of the reconstructed surface gradient image using the albedo independent PSM technique and the calculated height map using a FFT based method shown as a iso-height plot.

Face	X	Y	Z
A	0.054525	0.894764	-0.443198
B	-0.836804	-0.244101	-0.490076
C	0.513713	-0.449641	-0.730699

Table 4: Unit face normals of the milk package (markings in Fig. 11).

Fig. 11 shows orthogonal and perspective plots of the reconstructed milk package seen from two different view directions. Parts of the milk leave unreconstructed since only a 2.5-D model can be recovered from a single sensor position/direction. In Fig. 11 left an iso-height plot of a height map determined from the reconstructed 2.5-D model data is shown. For comparison, the gradient image has been transformed to a height map using a FFT-based method. An iso-height plot of this height map is shown in Fig. 10 right.

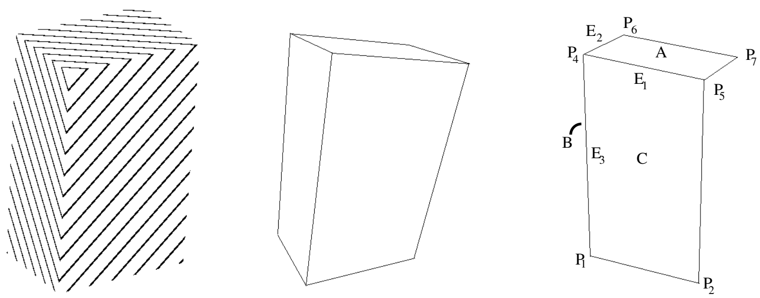

Figure 11: The calculated height map using our polyhedron reconstruction method shown as an iso-height plot and perspective plots of the 2.5-D reconstruction.

Tab. 4 and Tab. 5 summarize some data of the generated 2.5-D model. In Tab. 4 the determined unit surface normals are listed for the three visible faces. The values in Tab. 5 are given in the object coordinate system.

Vertex	X	Y	Z
P1	-0.186663	-0.763393	0.062641
P4	-0.133929	0.320592	-0.567323
P5	0.504743	0.461217	-0.204844
P6	-0.391741	0.508092	-0.220501

Table 5: 3D-points of four visible vertices (markings in Fig. 11).

The angles between the faces A, B and C are: angle(A,B) = 92.7, angle(A,C) = 92.9, angle(B,C) = 87.8. From the vertex data the ratios of the edge lengths can be calculated: E3:E1 = 1.6783, E3:E2 = 2.6638, E1:E2 = 1.5872. The real edge ratios are: E3:E1 = 1.7579, E3:E2 = 2.6508, E1:E2 = 1.5079. Using edge E2 as the reference edge the calculated lengths differ by 5mm (E1) resp. 1mm (E3).

Fig. 12 presents a part of our data base objects. The condition of the analyzed packages, concerning deformation and dirt is adequate. It is reasonable for the design of such a recognition system to expect large rates of incoming packages. The shading based reconstruction algorithm is able to process colored objects with a similar accuracy as uniform objects as shown in Sect. 3. As expected, the algorithm is not able to reconstruct dark parts of the package imprints. We process only 29 objects, since at this time the system doesn't contain a tool for the automatic data base generation. The system is designed to use the adaptive indexing of the geometric features to reduce the number of models to compare in the color indexing and the direct matching. In the described data base this preprocessing would shrink the remaining amount of data base objects too strong. With a larger data base, about 30 geometrically similar objects are the expected number of objects to be processed by the color indexing and the direct matching. Results of the adaptive indexing and learning algorithm are shown in [2].

To proof the capability of the color indexing and the direct matching we compared all data base objects with the segmented scene object faces. Tab. 6 and Tab. 7 shows results of the different color models and dirty and partially deformed objects. The results are performed by a combination of the color indexing and the direct image matching. The experiments using the $I_1I_2I_3$ color space produce the highest recognition rate. This is caused by the missing ability of the HSI color space to accuratly represent objects with dark colors. In this case the hue channel shows strong variations in the representation of homogenous regions. In the final version of our system we will use the direct matching of the imprints to verify the three best matching objects which are extracted by the color indexing. The direct matching shows a accuracy between 90%-95% to find the corresponding object, if we match objects with our total data base. Since the algorithm is very time consuming, we will only use it to for the mentioned verification. In the contrary to the color indexing, there is no possibility to integrate a precompiled optimization in the direct matching. The number of necessary operations depepends on the image size, the used resolution of the image analysis, the window size of the adaptations algorithm and the used iteration rate of the adaptation.

color space	$I_1I_2I_3$	RGB	HSI
recognized	72%	69%	62%
among 3 most similar objects	89%	93%	86%
not recognized	11%	7%	14%

Table 6: Comparison of three different color spaces used in the color indexing.

objects	data base	partially deformed and dirty objects
recognized	75%	71%
among 3 most similar objects	93%	86%
not recognized	7%	14%

Table 7: Weighted combination of the color indexing and the direct matching.

9 CONCLUSION

It has been shown, that a shading based shape recovery method (Photometric Stereo) leads to sufficient 2.5-D models for application in an elaborate object recognition system. Particularly this is appropriate, if the derived 3-D shape information is combined with an approximated line drawing. The line drawing is necessary if occlusions have to be taken into consideration. We have shown, that good results are reached when combining PSM, the ESNI model, the adaptive indexinig of geometric features, the color indexing and the direct matching. In the future, the shape recovery part of the project will be focused on making use of shadow information and mutual illuminations. The matching has to be optimized for the application in an industrial environment with the demand to process a high number of pieces.

10 REFERENCES

[1] G. Bellaire, Feature-Based Computation of Hierarchical Aspect-Graphs, *Machine GRAPHICS & VISION*, Vol. 2, no 2, pp. 105-122, 1993.

[2] G. Bellaire, K. Schlüns, A. Mitritz, K. Gewinner, "Adaptive Matching Using Object Models Generated From Photometric Stereo Images", *Proc. of 8th ICIAP*, San Remo, Italy, pp. 293-298, 1995.

[3] G. Bellaire, M. Lübbe, "Adaptive Hierarchical Indexing and Constrained Localization: Matching Characteristic Views", *Proc. of the ICSC'95*, Hong-Kong, pp. 241-249, 1995.

[4] P. Flynn, A.K. Jain,"CAD-Based Computer Vision: From CAD Models to Relational Graphs", *IEEE Trans. on PAMI*, Vol 13, No. 2, Feb, pp 114-132, 1991.

[5] B.K.P. Horn: *Robot Vision,* The MIT Press, Cambridge, 1986..

[6] K. Ikeuchi, K.S. Hong, "Determining Linear Shape Change: Toward Automatic Generation of Object Recognition Programs", *CVGIP*, Vol. 53, No. 2, pp. 154 - 170, 1991.

[7] G.J. Klinker, S.A. Shafer and T. Kanade, "A Physical Approach to Color Image Understanding", *IJCV*, 4, pp. 7-38, 1990.

[8] H. Lu, L. G. Shapiro, "Model-Based Vision Using Relational Summaries", *SPIE Vol 1095, Applications of Artificial Intellience VII*, pp. 662 - 675, 1989.

[9] J.H. Park, T.G. Chang, J.S. Choi, "Three-Dimensional Object Representation and Recognition Based on Surface Normal Images", *PR*, 26, No. 6, pp. 913-921, 1993.

[10] K. Schlüns: "Photometric Stereo for Non-Lambertian Surfaces Using Color Information", *Proc. 5th Int. Conf. on Computer Analysis of Images and Patterns*, Budapest, Hungary, pp. 444-451, Sept.1993.

[11] K. Schlüns and M. Teschner, "Fast Separation of Reflection Components and its Application in 3D Shape Recovery", *Proc. 3rd Color Imaging Conference*, Scottsdale, Arizona, pp. 48-51, 1995.

[12] S.A. Shafer, "Using Color to Separate Reflection Components", *COLOR research and application*, 10(4), Winter 1985, S. 210-218.

[13] MJ. Swain, D. Ballard, "Color indexing", *IJCV*, Vol. 7, No. 1, pp. 11-32, 1992.

[14] R. J. Woodham, "Photometric Method for Determining Surface Orientations from Multiple Images", *Optical Engineering*, Vol. 19, No. 1, pp. 139-144, 1980.

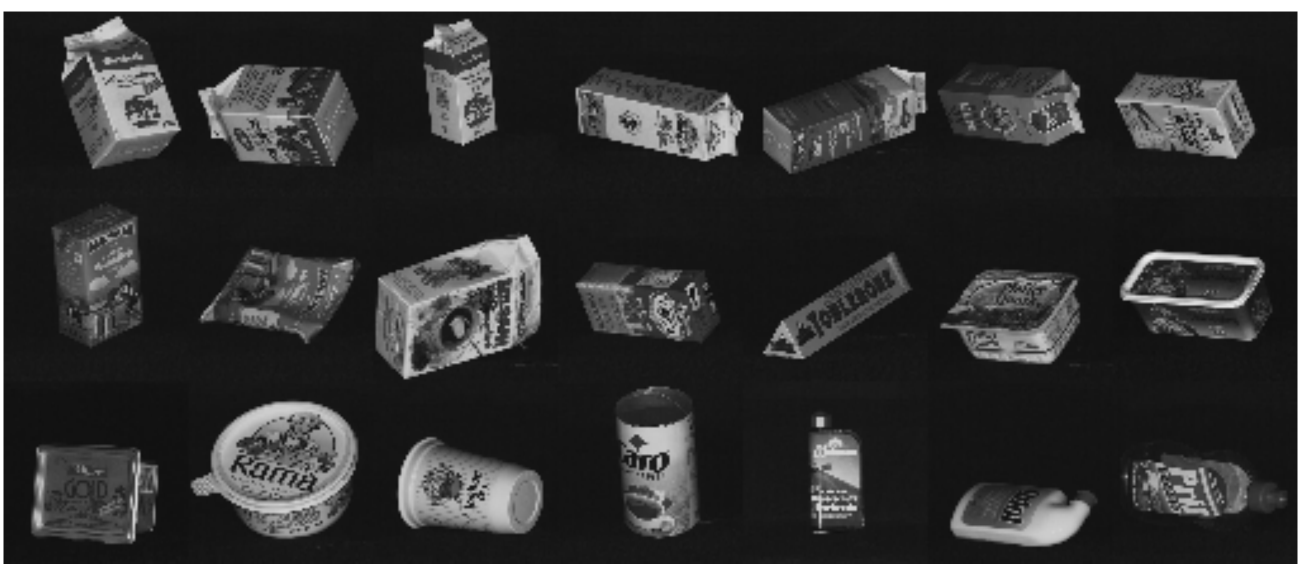

Figure 12: Data base objects that are processed in the experiments.