The Color Adjacency Graph Representation of Multi-Coloured Objects

J. Matas, R. Marik, and J. Kittler
Department of Electronic & Electrical Engineering,
University of Surrey, Guildford, GU2 5XH, UK

Technical Report VSSP-TR-1/95

Abstract

A new representation for objects with multiple colours - the colour adjacency graph (CAG) - is proposed. Each node of the CAG represents a single chromatic component of the image defined as a set of pixels forming a unimodal cluster in the chromatic scattergram. Edges encode information about adjacency of colour components and their reflectance ratio.

The CAG is related to both the histogram and region adjacency graph representations. It is shown to be preserving and combining the best features of these two approaches while avoiding their drawbacks. The proposed approach is tested on a range of difficult object recognition and localisation problems involving complex imagery of non rigid 3D objects under varied viewing conditions with excellent results.

1 Introduction

Consider the four images depicted in figures 1a-d. Very few observers will doubt (at least from viewing the colour originals) [1] that figures 1a-d depict the same object, regardless of the fact that different parts (in case of 1b non-overlapping parts) are imaged at significantly different scales. More importantly, the imaged object can be readily localised in the cluttered scenes of figs. 1e and 1f using only description of the object derived from fig. 1a (or 1b-d).

The problem of recognition of objects with complex colour structure has recently received increased attention [SB90] [Wix90] [Fat92] [SS94]. Efficient methods based on colour histograms produced impressive recognition rates [Swa90], but only if the object of interest had already been localised allowing the histogram to be computed over an area largely covered by the object. Furthermore, performance of colour histogram matching deteriorates rapidly when illumination intensity changes [SB90] [Fin92]; the same problem arises from changes in relative position with respect to the light source. The bin counts of a normalised colour histogram are proportional to ratios of projected areas, which are only invariant under affine transformation of a plane. Consider a cube with pairs of opposite sides coloured blue, red and green. *Any* proportion of red, blue and green may appear in the histogram depending on the viewpoint. It is therefore not surprising that almost all reported experiments on colour histograms either process images of effectively planar scenes or are entirely in the 2D domain (searching image databases [SS94]); in scenes with 3D structure, successful matching has been demonstrated only for images taken from slightly different viewpoints under constant lighting conditions [Swa90].

An alternative approach, based on matching of attributed region adjacency graphs (RAGs), was pursued by Fathima [Fat92]. Recognition is implemented by subgraph matching of RAGs describing the scene and the object model(s). The method works well for simple scenes, but the computational complexity of subgraph matching together with the size of RAGs for common scenes and objects severely limits applicability of the the approach (see figures 2, 6). A further complication of the problem stems from the need to store and match RAGs corresponding to a (often large) number of aspects.

In this paper we present a new representation for objects with multiple colours - the *colour adjacency graph* (CAG) . Each node of the CAG represents a single colour (chromatic) component of the image (or object); a component is defined as a set of pixels forming a unimodal cluster in the chromatic histogram. Edges of the CAG contain information about adjacency of colour components. Each edge has as an attribute the reflectance ratio, a photometric invariant proposed by Nayar and Bolle [NB93b]. The CAG is related to both the histogram and RAG representations. Nodes of CAG correspond to *modes* (peaks) of the histogram. Location of the maxima, unlike the associated bin *counts* are extremely stable under viewpoint change and are robust to occlusion. Surprisingly, CAGs of many objects man-made and natural objects have only one or two aspects! In other words, the number of colored regions and their relative size may change with the viewpoint, self-occlusion and non-rigid transformation; but the set of visible colours and their adjacency relationship remains stable. This is exactly the case for the ICPR proceedings in figures 1a-f; but think of a tiger in motion, a bush of roses, the books on your shelf or

[1] Colour originals of images shown in the paper are available by anonymous ftp to ftp.ee.surrey.ac.uk, directory /pub/vision/images/cag-iccv95

E-mail contact: g.matas@ee.surrey.ac.uk

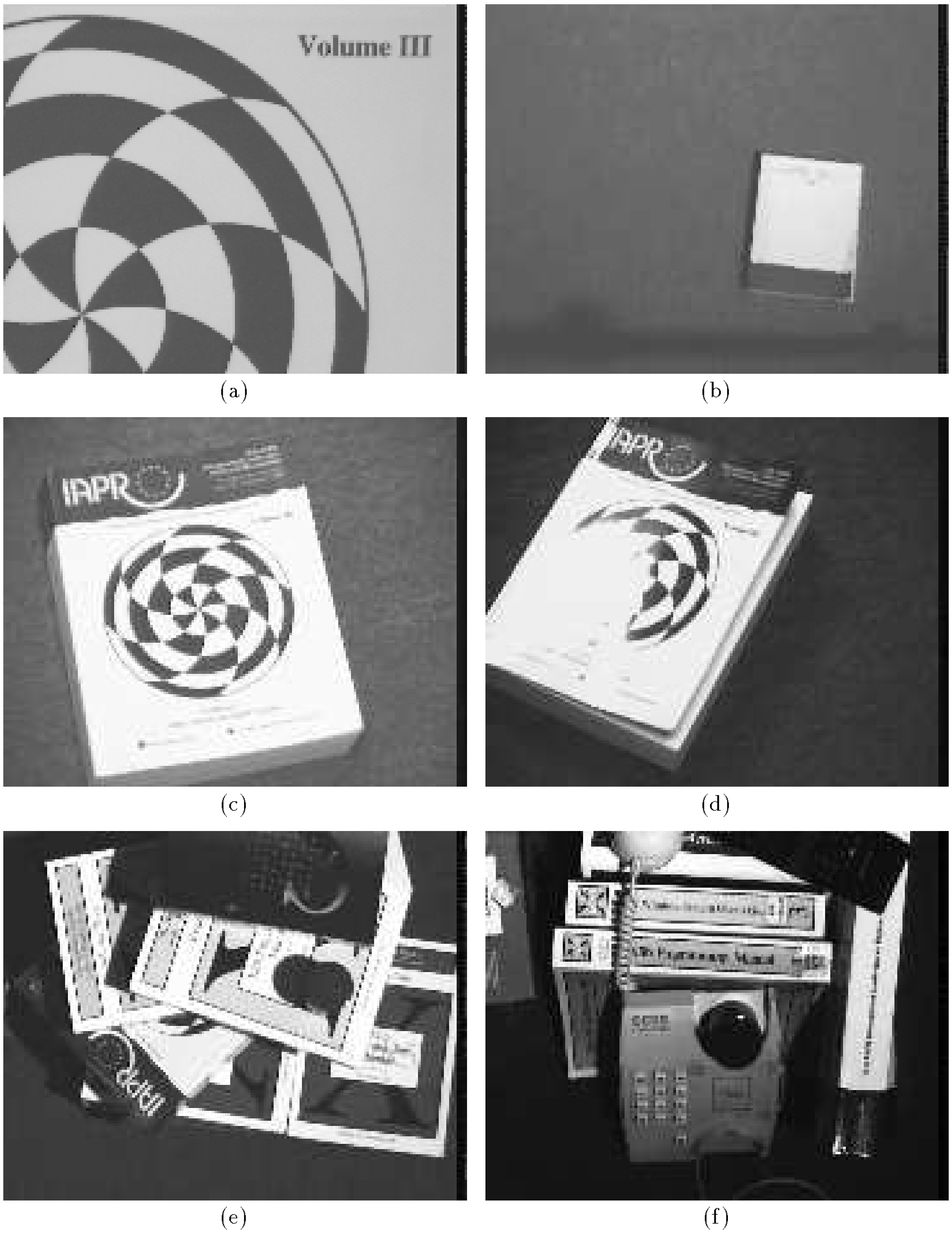


Figure 1: ICPR proceeding. (a) Detailed view, front. (b) Distant view, back. (c) Full view, front. (d) Full frontal view with a specularity. (e)(f) Scenes with occluded views of the proceedings. Colour and adjacency graphs of (a)-(d) are depicted in fig. 2. Recognition results for scene (f) are shown in fig. 9

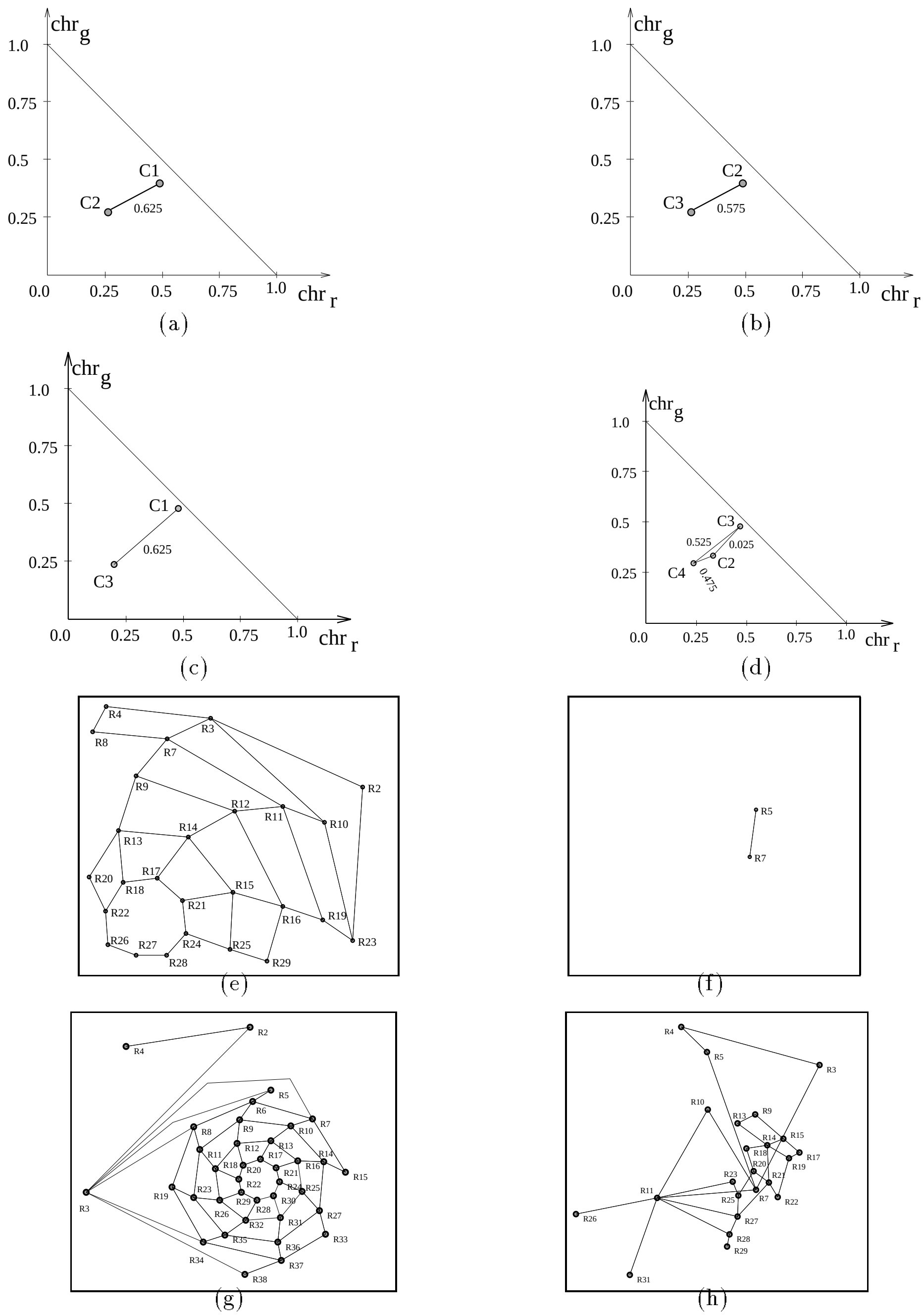


Figure 2: ICPR proceedings. Colour (a-d) and region (e-h) adjacency graphs for images shown in fig.1a-d. The difference in stability of the CAG and RAG representation is striking. The CAGs are embedded in the chromatic plane to show chromaticity values associated with nodes. RAGs are embedded in image plane with nodes positioned at region barycenters. CAG edges are attributed with reflectance ratios. Subimages (d) and (h) show the influence of a large specularity on the two representations. The effect is identical to on occlusion by a white object. If the CAG of (d) is matched against any of (a)(b)(c) node C2 of (d) is left unmatched.

even yourself, as you are dressed now! But in general object CAG does change under the transformations discussed above. We therefore define an object CAG as a *union* of all observed image CAGs for the object. After sufficient number of observations, the set of nodes in the union should be equivalent to the set of colours on the surface of the object and the set of edges identical to the set of all colour boundaries.

The problem of localisation and recognition of objects can be posed as a search for subgraphs of object CAG(s) and the image CAG. The CAG subgraph matching does not suffer from the complexity problems of RAG matching for two reasons. First, edge matches, requiring agreement of five attributes (two chromaticities and the reflectance ratio), are rarely ambiguous. Moreover, in most of our experiments, the complexity of CAGs was at least an order of magnitude smaller then that of RAGs (see eg. fig. 6).

The rest of the paper is organised as follows. Section 2 presents the adopted reflection model. We show that, for this common reflection model, the attributes of the CAG are invariant to viewing and illumination geometry as well as illumination intensity. Section 3 describes in detail how the CAG is built. In subsection 3.1 we discuss the graph-theoretical non-parametric clustering algorithm used for separation of modes in the chromatic histogram. Detection of color adjacency and robust computation of the reflectance ratio is presented in section 3.2. Matching of colour adjacency graphs is described in section 4. Recognition experiments are reported in 5. Results are summarised in section 6.

2 Reflection Model

In [Hea89b] Healey showed that reflectance S of an opaque material can be accurately approximated as:

$$S(g,\lambda) = \begin{cases} m_i(g)C_i(\lambda) & \text{metals} \\ m_b(g)C_b(\lambda) + m_i(g)C_i(\lambda) & \text{dielectrics} \end{cases} \tag{1}$$

where g and λ indicate dependence on photometric angles and wavelength respectively. $C(\lambda)$ and $m(g)$ are, in general, arbitrary functions. Subscripts i and b denote terms associated with interface (specular) and body reflection. Following [Hea89a] we simplify the second part of equation (1) (equivalent to the dichromatic reflection model [Sha84]) by assuming that effects of interface reflection for (inhomogeneous) dielectrics are negligible over most pixels in the image. Moreover, high intensity of specular points often saturates the sensor which significantly reduces possibilities of colour analysis. The effect of a large specularity on the CAG representation is shown in figs. 1d and 2d.

Introducing the simplification to (1) we obtain

$$S(g,\lambda) = m(g)C(\lambda) \tag{2}$$

Sensor response of a standard imaging device is well modelled by a spectral integration process:

$$\begin{aligned} p_k^X &= \int_{\lambda_1}^{\lambda_2} \rho_k(\lambda) S(g,\lambda) E(\lambda) d\lambda, \qquad k = 1, \cdots, n && (3) \\ &= m(g) \int_{\lambda_1}^{\lambda_2} \rho_k(\lambda) C(\lambda) E(\lambda) d\lambda && (4) \end{aligned}$$

where p_k^X is the response of the k-th sensor at location X of the sensor array, $E(\lambda)$ is the spectral power distribution (SPD) of the light incident on the surface patch that is projected on pixel X, $\rho_k(\lambda)$ is the responsivity function of the k-th sensor and n is the number of sensors at X. In our experiments we used standard RGB equipment; we will therefore assume $n = 3$ and $\bar{p} = (p_1, p_2, p_3) = (r, g, b)$, where for brevity we drop the explicit notation of the output dependence on sensor location X. From (4) it follows that measured values corresponding to a single material illuminated by a source with SPD $E(\lambda)$ of a strength controlled by a constant (ie. $kE(\lambda)$) will lie on a line passing through the origin.

We elected to represent the linear cluster by its intersection with the chromatic plane, ie. the plane passing through $(0,0,1)$, $(0,1,0)$, $(0,0,1)$ in the RGB space. Computing the intersection and substituting from eq. (4) we obtain:

$$\begin{aligned} ch_r &= p_1/(p_1+p_2+p_3) = r/(r+b+g) && (5) \\ ch_g &= p_2/(p_1+p_2+p_3) = g/(r+b+g) && (6) \\ &= \frac{\int_{\lambda_2}^{\lambda_2} \rho_1(\lambda)C(\lambda)E(\lambda)d\lambda}{\sum_{k=1}^{3}\int_{\lambda_1}^{\lambda_2} \rho_k(\lambda)C(\lambda)E(\lambda)d\lambda} && (7) \end{aligned}$$

It is clear from eq.(7) that for the chosen reflection model chromatic values are independent of viewing geometry and illumination intensity. Sensor noise, variations in material composition, specularities and interreflections will cause the observed values to deviate from the line. We treat the deviations as random noise. We do not model the error distribution of the deviations, but only assume that smaller errors are more likely than larger ones. Under this assumption, the noisy linear cluster corresponding to a material of a certain reflectance will form a unimodal cluster in the 2-dimensional chromatic space and the detection of such clusters can be rephrased as the problem of unsupervised clustering of unimodal distributions [Kit76].

In the paper we do not address the problem of changing illumination SPD. We avoid the issue with a standard disclaimer, ie. 'Input image is assumed to be processed by a colour constancy algorithm (eg. [MB86])'. In practise, we performed all matching experiments on CAGs from images illuminated with light of identical SPD. We believe that problems related to changes in illumination do not reduce the usefulness of the CAG representation. On the contrary - CAG is build in an illumination independent way and may even provide useful input (eg. correspondence) to colour constancy algorithms.

3 Building the Color Adjacency Graph

In the recognition experiments described in section 5 a CAG description of a scene is matched against CAGs representing objects. The form of object and images CAG is identical but minor differences exist in the building stage. In the model acquisition stage, image CAG is transformed into a single view object CAG by removing nodes corresponding to background. Next, if more than one view is needed to obtain a full object CAG, a CAG-union operation must be performed. The union operation was implemented as CAG matching (section 4) followed by insertion of unmatched nodes.

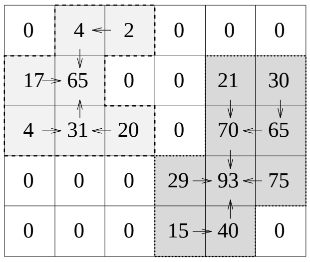


Figure 3: Graph-theoretical clustering in 2D space. Arrows symbolise links to bins with the largest count in local neighbourhood. Different shading highlights the two clusters of the example. Although the example used 4-neighbourhood, in this particular case (and generally with well separated clusters) the result is fairly insensitive to the size of searched neighbourhood. In this example, identical clustering is obtained with 8 and 12 neighbourhoods.

Color Label	Chromaticity		Cluster Maximum	Cluster Size
	r	g		
1	0.385	0.375	10895	53245
2	0.205	0.445	8173	17810
3	0.335	0.285	6375	20067
4	0.355	0.315	2555	8703
5	0.265	0.395	106	403
6	0.295	0.365	61	351

Table 1: Results of the graph-theoretical clustering on the 'pyjamaTop' image (fig. 4a). The clusters are detected in the order defined by the bin count of the maximal bin (peak size). Chromaticity associated with the maximal bin (second and third columns) attributes the corresponding CAG node. The total number of pixels in each cluster is shown in the last column. After boundary pixel removal, no pixels are left in clusters 5 and 6. Knowing that cluster 1 originates from the background, only clusters 2,3 and 4 are inserted in the object CAG (see fig. 4e)

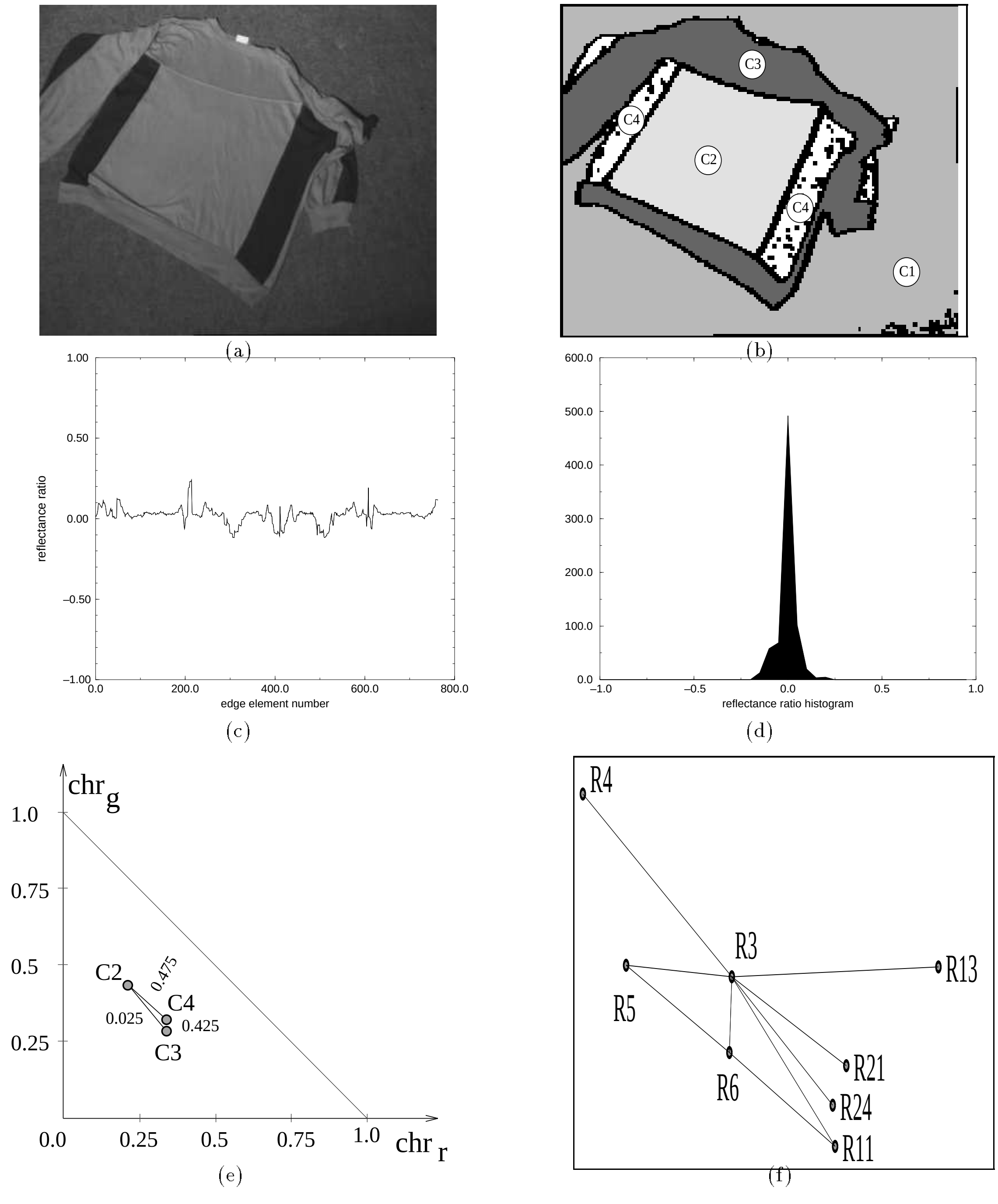


Figure 4: Building the CAG. Acquisition of the 'pyjamaTop' model. (a) Original image. (b) Results of clustering after boundary pixel removal. Pixels are labelled according to cluster membership. (c) Reflectance ratios along the boundary between components C2 and C3. (d) Histogram of (c). (e) CAG. The largest component, C1, is known to belong to background and therefore not included in the object CAG. (f) RAG. The very low contrast between components C2 and C3 in (a) is consistent with the reflectance ratio distribution around 0.

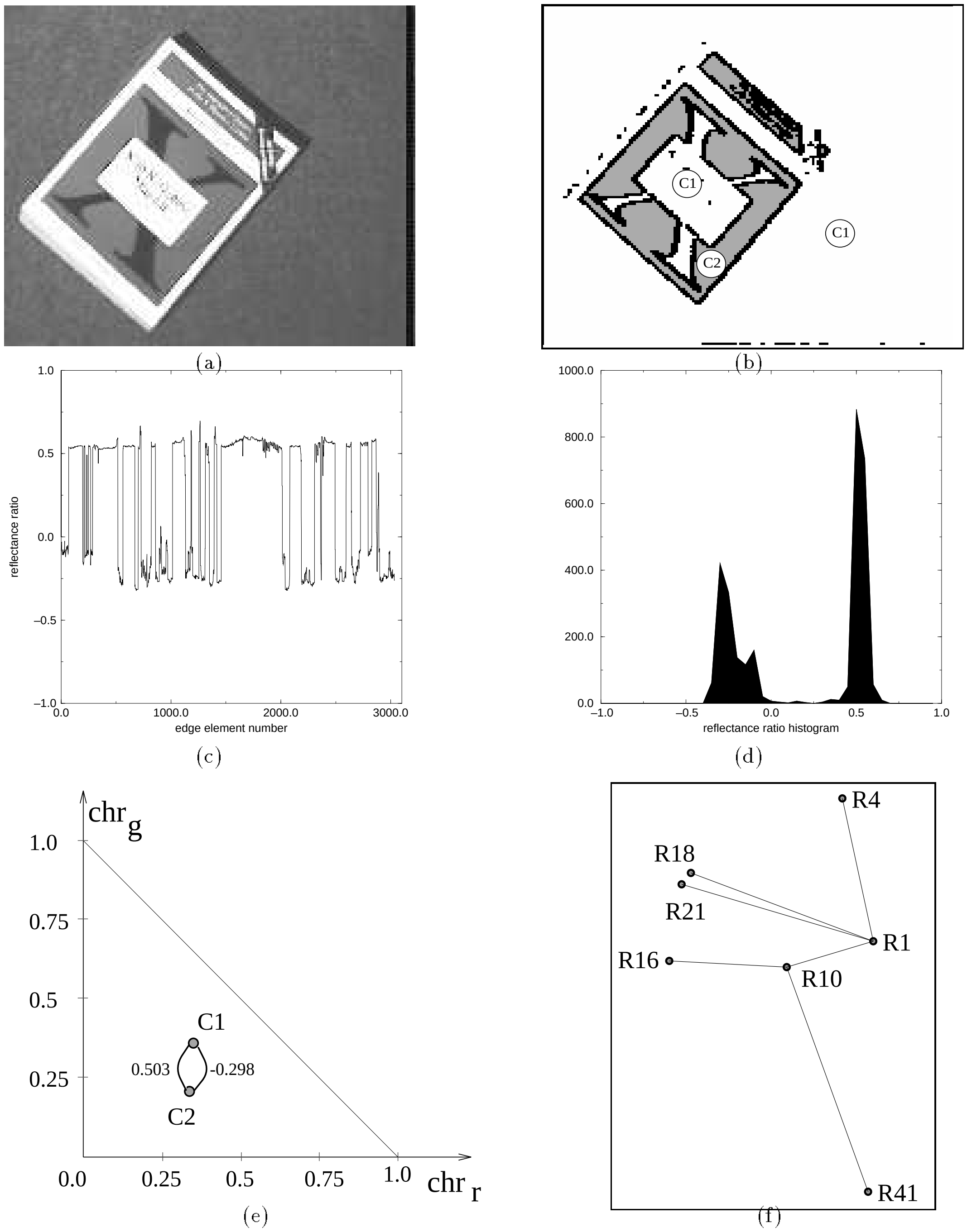


Figure 5: Building the CAG. Acquisition of the 'manual3' model. (a) Original image. (b) Results of clustering after boundary pixel removal. Pixels are labelled according to cluster membership. (c) Reflectance ratios along boundary between components C1 and C2. (d) Histogram of (c). (e) CAG. (f) RAG. The two reflectance ratio modes are represented by two edges between components C1 and C2 in (e). Modes of reflectance ratios are detected by a 1D form of graph-theoretical clustering. The small peak close to 0.0 in (d) does not define a mode, because large values to the left are in its neighbourhood (neighbourhood diameter 0.15).

3.1 Nodes

In section 2 we showed that, under the assumed reflection and noise model, *r,g,b* values of pixels imaging the same material form a unimodal cluster in the chromatic plane. We call this set of pixels a colour component. Each node of the CAG represents a *colour component*. The problem of unimodal cluster separation has received significant attention in the pattern recognition community [KNF76][KB90]. We adopted the Graph-theoretical clustering method of Koontz and Fukunaga. The method can be outlined as follows (for details and comparison with other approaches see [Fuk90]).

Algorithm 1: **Graph-theoretical clustering (in chromatic plane)**

1. Compute chromaticity histogram. In each bin, maintain a list of pixels voting for the bin.
2. For each bin, find the bin with maximal count in a given neighbourhood. Store a link to this bin.
3. Such links form a forest, with a root of each tree in a local maximum. The set of pixels voting for bins in a single tree form a unimodal cluster.

The graph-theoretical clustering algorithm requires two parameters: the size of a histogram bin and the size of the neighbourhood considered when searching for maximal bins in local neighbourhood. Methods for automatic selection of these parameters are given in [KB90].

The result of the graph-theoretical clustering for a simple example is shown in figure 3. The arrows symbolise links to maximal bins in the local neighbourhood (in this case 4-neighbourhood). Shaded areas correspond to clusters detected by the method.

In our implementation we compute the links to maximal neighbourhood bins in a single sweep. In the same pass we insert indices of all local maxima in a list. We sort the list according to the bin count. Next, bins belonging to the cluster associated with the local maximum currently at the top of the list are obtained by traversing the tree rooted at the local maximum. Note that this implementation first outputs the cluster with a mode in the global maximum, other salient clusters follow. The algorithm therefore need not (and typically does not) completely segment the feature space. The process can be terminated if eg. a cluster with interesting properties was found or a sufficiently large part of pixels have been analysed. The process can also be easily tuned to serve a top-down process by eg. considering maxima only from a selected part of the chromatic space, corresponding to a specific object.

Table 1 summarises performance of the graph-theoretical clustering for the image shown in fig. 4(a). The chromaticity of the maximal bin in each cluster is used as attribute of the node. Other attributes based on more complex description of the cluster could be trivially added. In a post-processing step, pixels lying on colour component boundaries are removed by erosion. In general, boundary pixel colour is a mixture of colour inside and outside the colour component and is therefore unreliable. The backprojected colour components after erosion for image 4a are shown in 4b and for image 5a in 5b.

3.2 Edges

Edges of the CAG represent adjacency of colour components. We proceed as follows. Neighbourhood of each boundary pixel in the eroded labelled image (figs. 4b, 5b) is searched for the nearest pixel belonging to a different colour component. If such pixel is found, the pair of pixels is inserted in the list of boundary pixels for the two colour components involved. An edge is inserted into the CAG between the nodes representing colour components if the size of their boundary list is larger than a small threshold (default set to 5).

Nayar and Bolle [NB93a] show both experimentally and theoretically that, under a fairly general reflection conditions, intensity ratios on smooth surfaces are photometrically invariant end equivalent to 'reflectance ratios'. To avoid the mathematical inconveniences caused by the fact that an intensity ratio can attain values in the range $\langle 0, \infty)$ we adopt the convention of [NB93a] and define the ratio as:

$$r_I = \frac{I_1 - I_2}{I_1 + I_2} \tag{8}$$

which maps the ratio conveniently into the interval $\langle -1, 1 \rangle$, with the sign expressing the 'brighter than' relationship. We analyse the intensity ratios in the list of boundary pixels using the 1D version of the graph-theoretical clustering mode. If a single mode is found, the ratio corresponding to the maximal bin attributes the edge. If more modes are found, the edge is split. As in the case of nodes, a more precise description of the ratio histogram could lead to a large number of edge attributes. Examples of a set of reflectance ratios and their histograms are given in figs. 4c, 4d. If the smoothness constraint on the boundary is violated, then different ratios will be observed for the same pair of colour components and inserted in the CAG. Matching of such colour components will not be affected by the ratios, as any possible ratio will be represented in the model.

The benefit of using intensity ratios is clearly seen in the example shown in fig. 5. A purely chromatic (or hue and saturation) approach would describe the object colour structure with four numbers, namely the chromaticity (hue-saturation) of gray (the black and white part) and of blue. In our approach, the same object is described by six numbers, which greatly reduces a random mismatch.

4 CAG Matching

If a scene contains an object whose CAG is known, then a subgraph of the scene CAG will match a subgraph of the model CAG corresponding to the visible part of the object. Recognition and localisation of objects using CAG representation can be therefore accomplished by subgraph matching. The work presented in the paper focusses on the usefulness of the CAG representation and not on graph matching per se, we therefore implemented a simple algorithm performing effectively a best-first search with no backtracking. The algorithm provides a viable alternative to the more sophisticated methods (see eg. [CKP] or references therein) because of a very low probability of mismatch between CAG edges. Our algorithm can be outlined as follows:

Algorithm 2: **Simple CAG matching**

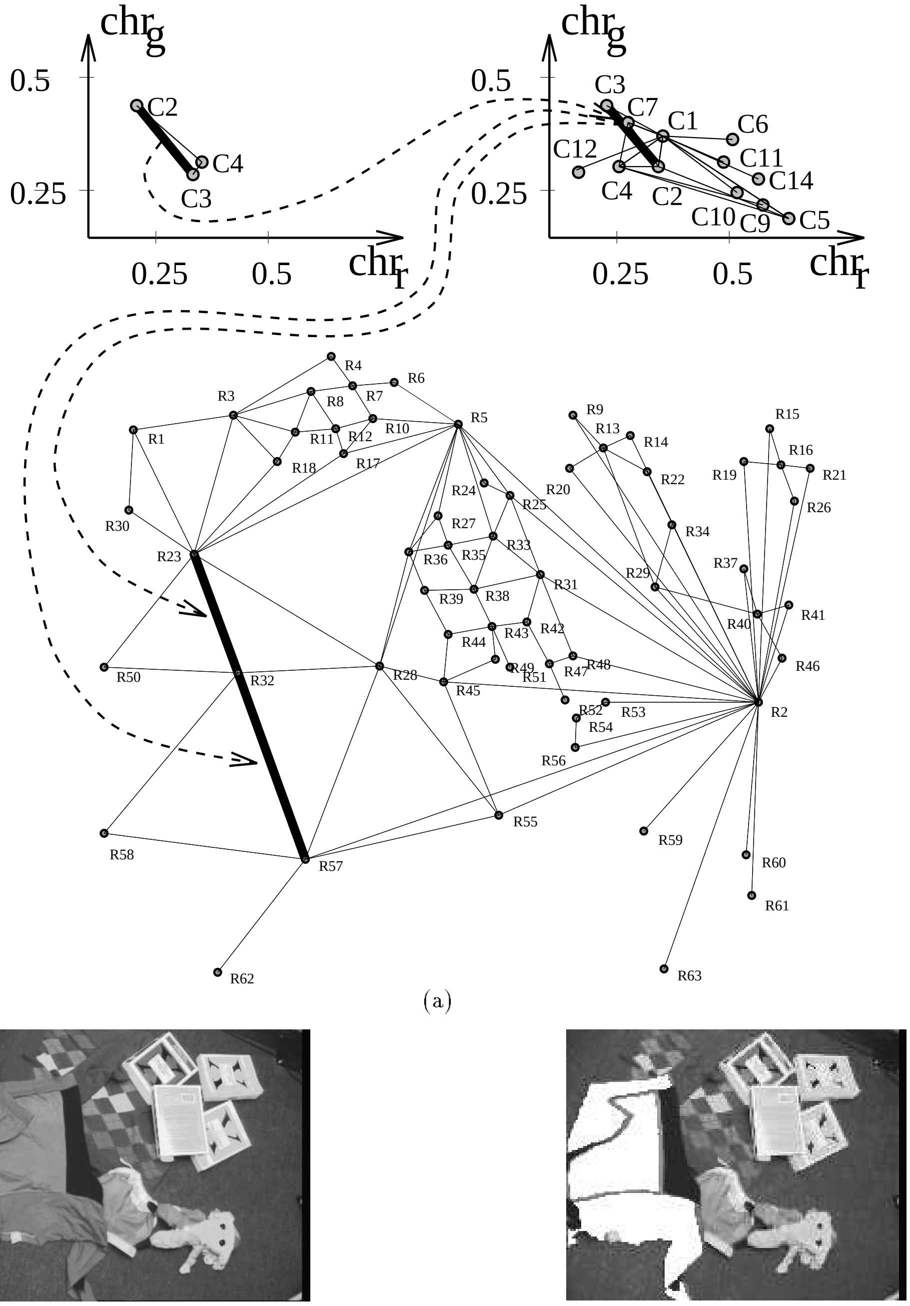


Figure 6: Matching in Experiment 3. (b) The scene. (a) The scene CAG is shown in the top-right corner. The scene RAG is depicted in the center. The top-left corner shows the 'pyjama' model CAG. During the recognition process edge C2,C3 of the 'pyjama' model CAG is matched to edge C3,C2 of the image CAG. The match is propagated in the RAG and backprojected into the image. In this particular case, three region are labelled. By coincidence, CAG edges between nodes with identical indices (C2, C3) are matched. This is due to the fact that the coloured patches of the 'pyjama' belong to the largest colour components of the image after the background; their relative size is reversed, as C2,C3 is matched to C3,C2.

1. Let $\mathcal{M}$ be the set of CAG edges of all models in the database.
   Let $\mathcal{I}$ be the set of edges of the image CAG.
   Let $\mathcal{P}$ be the set of all pairs (e_M, e_I), $e_M \in \mathcal{M}$, $e_I \in \mathcal{I}$
2. For all $(e_M, e_I) \in \mathcal{P}$ compute quality of the edge match.
3. Order pairs (e_M, e_I) according to match quality.
4. **While** best match quality higher then threshold
   if edge match is consistent with previous matches
   (a) label the image CAG edge
   (b) label the image regions corresponding to the matched edge

---

A weighted sum of differences in chromaticity and reflectance ratio was used as the criterion of edge match quality. Two tests of consistency are performed. A single model CAG edge can only match a single image CAG edge and vice versa. The second test checks the consistency of image interpretation. Before starting the matching process, image RAG is computed by connected component analysis of colour components. Once an image CAG edge is matched, the pair of regions mapping into the image CAG edge are labelled as belonging to the matching model CAG. Any match that would attempt to relabel a region is not accepted. An example of a pair of region labelled as a result of a single CAG edge match is shown in fig. 6.

Note that although the algorithm assumes that CAG *edges* unambiguously identify the model, no such assumption is made about *nodes*. It is common that a colour component (ie. the set of pixel corresponding to an image CAG node) with a common chromaticity, eg. white-gray-black, contains regions of more than one object. This is clearly illustrated by the example of the seven X manuals, which all have white and black patches with identical chromaticities. Finally we remark that although the CAG can represent an object of single colour by graph with a single node, such object cannot be matched be Algorithm 2.

5 Experiments

A number of experiments were performed to demonstrate the effectiveness of the proposed colour based object recognition method. The experiments were designed to illustrate the stability of the method under change of viewpoint and in the presence of interfering artifacts. Robustness of CAG and a comparison with the RAG has already been discussed in the paper and is demonstrated by the results shown in figs. 1 and 2.

Matching experiments were performed with a model database containing eleven objects (seven different manuals, a manikin, a pyjama, a sweater, and an ICPR proceedings) pictured in Figure 10. The object models were acquired as described in Section 3. The complexity of the CAGs varied from a two node graph for the ICPR proceedings to a 4 node graph for models 'manikin' and 'sweater'.

The sequence of images used in Experiment 1 is shown in the top row of Figure 7. In addition to a subset of known objects, each frame contains unknown objects whose model is not in the database. Thus, for instance, Figure 7(a) contains the background

Figure 7: Experiment 1. (a-f) Processed images. (g-l) Regions matched to model 'sweater' (see fig. 10(i)) are shown in white. Clustering results are visualised by setting to black the areas corresponding to pixels removed on the boundaries of colour components (see sect. 3.2). Successful recognition and localisation of the 'sweater' in images (a-e) is shown in (g-k). No false positive is reported in (l). Subimage (m) shows the answer to the query 'Any other known objects in (d)?'. No other objects are present, therefore nothing is highlighted. (n) Answer to the same query for (e). The 'pyjama' object is found and highlighted. (o-p)Answer to the same query, image (f). Both the 'manikin' and 'pyjama' objects are found. Black trousers are incorrectly recognised as part of the 'pyjama' (see text).

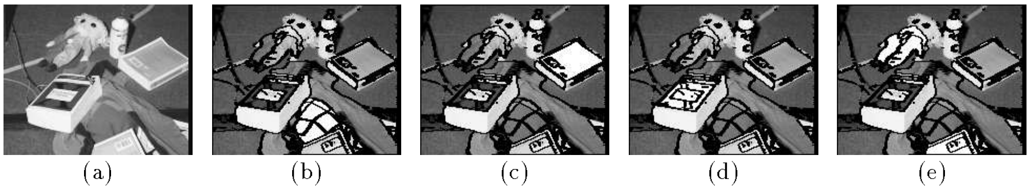
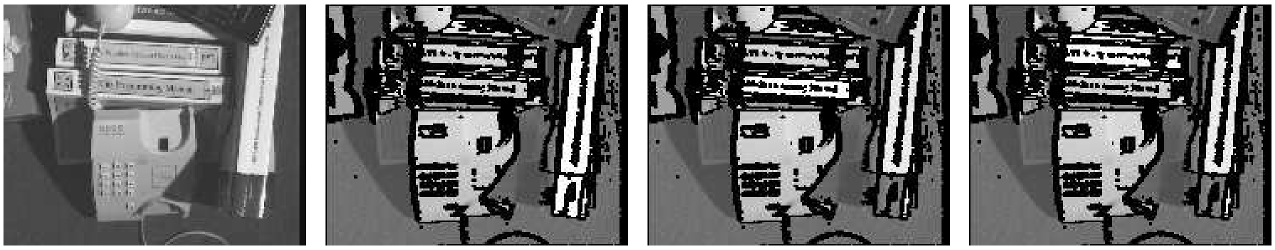
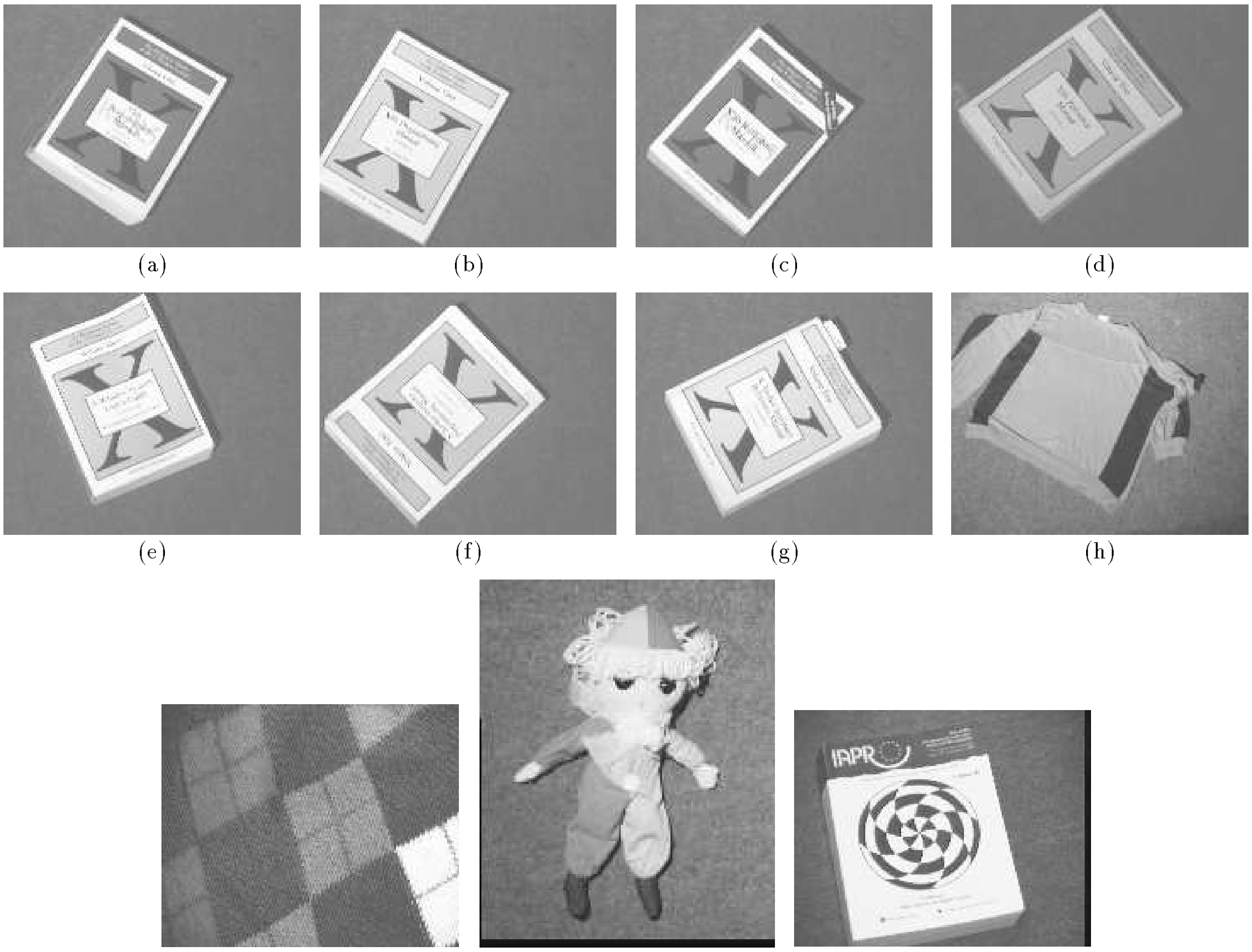


Figure 8: Experiment 2. (a) The scene. (b-e) Recognised objects: 'sweater' (b), 'manual6' (c), 'manual3' (d), 'manikin' (e). The bottle, which is not in the model database, is not matched .



(a) (b) (c) (d)

Figure 9: Experiment 3. (a) The scene. (b-d) Recognised objects: 'ICPR proceedings' (b), 'manual2' (c), 'manual5' (d).



(a) (b) (c) (d)

(e) (f) (g) (h)

(i) (j) (k)

Figure 10: The model database: (a-g) objects 'manual1' to 'manual7'. (h) 'pyjama' ('pyjamaTop') (i) 'sweater' (j) 'manikin' (k) 'ICPR proceedings'.

and Navid's head the chromatic properties of which are not known to the system. The unknown objects in Figure 7(d) also include Navid's stripy shirt which is also not included in the model database.

The second row of Figure 7 contains the results of processing in response to the combined query "Does the image contain the sweater and if so, where is it?". In principle, the result could have been obtained by trying to match a single model CAG to the image CAG. Instead, as in all matching experiments reported in this section, all the models were matched and results relevant to a particular query visualised. The pixels of the detected object are annotated in white. The boundary pixels between two adjacent colours are shown in black to allow the reader to see the outcome of clustering (segmentation). We can see from the top row of Figure 7 that the object of interest undergoes dramatic transformations due to changing viewpoint and due to its nonrigidity. Thus in (c) Navid's torso is rotated by 90 degrees as compared to the pose in (a). As the surface covered by the sweater is nonplanar, the areas of the colour patches on the sweater are distorted nonlinearly which would detrimentally affect both the histogram and RAG representations. In Figures 7 (d) and (e) the appearance of the sweater is changed dramatically as a result of its nonrigid nature. In (d) the sweater is thrown over an unknown object while in (e) the sweater is thrown over the pyjama. Figure 7 (b) just illustrates the presence of occlusion and in Figure 7 (f) the sweater is not present. We can see in Figures 7 (g)-(k) that the sweater is always correctly detected and segmented out. No false alarms are reported for Figure 7 (l).

The last row of Figure 7 presents the results of the query "What else is in the image and where?" Nothing is detected in Figure 7 (d) as shown in (m) where as the pyjama is found and segmented out in (e) as shown in (n). In Figure 7 (f) we found both the pyjama and the manikin. Note that the result contains a segmentation error. The trousers which are black and have the same colour as one of the constituent chromatic components of the pyjama. As these two regions happen to be adjacent to each other, the CAG is unable to discriminate between the black patch of the pyjama and the black trousers adjacent to other patches of the pyjama.

The processing of each image in the sequence takes about twenty seconds on SUN Sparc 10. It is dominated by the computation of the reflectance ratios (search for nearest neighbour) which is at the moment performed by a brute-force algorithm. Neither has any effort been invested in optimising the code for speed of processing. Even then the computation complexity compares favourably with the RAG approach which has been reported to require processing times up to several orders of magnitude longer.

The subsequent Experiments 2 and 3 were designed to illustrate a similar behaviour of the proposed methodology on different sets of images, shown in Figure 8 and 9 respectively. The same object model database was used in the experimentation. In each case the queried objects were detected and segmented out of the image correctly. The results demonstrate the robustness of the advocated method to shape and illumination changes, occlusion, and a widely changing viewpoint, including scale change.

6 Conclusions

A new representation for objects with multiple colours - the colour adjacency graph (CAG) - has been proposed. Each node of the CAG represents a single chromatic component of the image defined as a set of pixels forming a unimodal cluster in the chromatic scattergram. Edges encode information about adjacency of colour components and their reflectance ratio.

The CAG has been shown to be related to both the histogram and region adjacency graph representations. It preserves and combines the best features of these two approaches while avoiding their drawbacks. In particular, the nodes of the CAG correspond to peaks of the chromatic histogram. However the locations of the maxima, unlike the associated bin counts, are extremely stable under viewpoint change and are robust to occlusion. In comparison with the region adjacency graph, the CAG is significantly simpler, facilitating efficient matching. Moreover, its edge attributes (two chromaticities and reflectance ratio) are very rarely ambiguous.

It has been shown that the CAG can be generated very efficiently by means of theoretical graph clustering. Object recognition is achieved by subgraph matching of the image and model CAG's. Multiple instances of the same object or several distinct objects can be recognised and segmented out simultaneously.

The proposed CAG approach has been applied to a representative range of images containing rigid and nonrigid objects under varied viewing conditions, and subject to occlusion, change of scale, specular reflection distortion and 3D aspect change. The capability of the method to recognise and segment complex objects makes it a powerful tool of colour based object hypothesis generation and localisation.

Acknowledgements

The work was carried under ESPRIT project 7108 *Vision as process* and the BRITE project BREU-CT91-0496 *Automated Visual Inspection Applied to Stone Process.* We acknowledge Navid Fatemi-Ghomi for his help with Experiment 1.

References

[CKP] W.J. Christmas, J. Kittler, and M. Petrou. Matching in computer vision using probabilistic relaxation. Accepted for publication by *IEEE Trans. on Pattern Analysis and Machine Intelligence.*

[Fat92] S.T. Fathima. Data and model-driven selection using color regions. In *Image Understanding Workshop (San Diego, CA, Jan 26-29, 1992)*, pages 705–716, San Mateo, CA, 1992. Defense Advanced Research Projects Agency, Morgan Kaufmann.

[Fin92] Graham D. Finlayson. Colour object recognition. Master's thesis, Simon Fraser U, 1992.

[Fuk90] Fukunaga90. *Introduction to Statististical Pattern Recognition.* Academic Press, 1990.

[Hea89a] G. Healey. A parallel color algorithm for segmenting images of 3-d scenes. In *Image Understanding Workshop (Palo Alto, CA, May 23-26, 1989)*, pages 1038–1042, San Mateo, CA, 1989. Defense Advanced Research Projects Agency, Morgan Kaufmann.

[Hea89b] G. Healey. Using color for geometry-insensitive segmentation. *Journal of the Optical Society of America*, 6:86–103, June 1989.

[KB90] A. Khotanzad and A. Bouarfa. Image segmantation by a parallel , non-parametric histogram based clustering algorithm. *Pattern Recognition*, pages 961–973, 1990.

[Kit76] J. Kittler. A locally sensitive method for cluster analysis. *Pattern Recognition*, 8:23–33, 1976.

[KNF76] W.L.G. Koontz, P. M. Narendra, and K. Fukunaga. A graph-theoretic approach to nonparametric clustring. *Trasn. IEEE Computers*, C-25:936–944, 1976.

[MB86] L.T. Maloney and B.A.Wandell. Color constancy: a method for recovering surface spectral reflectance. *J. Opt. Soc. Am.*, 3:29–33, January 1986.

[NB93a] Shree K. Nayar and Ruud M. Bolle. Computing reflectance ratios from and image. *Pattern Recognition*, 26:1529–1542, 1993.

[NB93b] Shree K. Nayar and Ruud M. Bolle. Reflectance ratio: A photometric invariant for object recognition. In *Fourth International Conference on Computer Vision (Berlin, Germany, May 11-14, 1993)*, Washington, DC., 1993. IEEE Computer Society Press.

[SB90] M.J. Swain and D.H. Ballard. Indexing via color histograms. In *Third International Conference on Computer Vision (Osaka, JAPAN, December 4–7, 1990)*, pages 390–393, Washington, DC., 1990. IEEE Computer Society Press.

[Sha84] S.A. Shafer. Using color to separate reflection components. Technical report, University of Rochester, 1984.

[SS94] M. Stricker and M. Swain. The capacity fo color histogram indexing. In *CVPR'94 (IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Seattle, 1994)*, pages 704–708, Washington, DC., Jun 1994. Computer Society Press.

[Swa90] M.J. Swain. *Color Indexing*. PhD thesis, University of Rochester, 1990. prize (\$8.00).

[Wix90] L.E. Wixson. Real-time qualitative detection of multi-colored objects for object search. In *Image Understanding Workshop (Pittsburgh, Penn., Sep 11-13 1990)*, pages 631–638, San Mateo, CA, 1990. Defense Advanced Research Projects Agency, Morgan Kaufmann.