Probabilistic Indexing for Object Recognition

Clark F. Olson

***Abstract*—Recent papers have indicated that indexing is a promising approach to fast model-based object recognition because it allows most of the possible matches between sets of image features and sets of model features to be quickly eliminated from consideration. This correspondence describes a system that is capable of indexing using sets of three points undergoing three-dimensional transformations and projection by taking advantage of the probabilistic peaking effect. To be able to index using sets of three points, we must allow false negatives. These are overcome by ensuring that we examine several correct hypotheses. The use of these techniques to speed up the alignment method is described. Results are given on real and synthetic data.**

***Index Terms*—Object recognition, indexing, probabilistic algorithms, probabilistic peaking effect, alignment method.**

I. Introduction

Feature set indexing is a tool that is useful in object recognition algorithms. These indexing methods determine which sets of model features could have projected to various sets of image features, eliminating the need to consider other sets of model features as possible matches. This correspondence discusses indexing techniques for the problem of recognizing three-dimensional objects represented by feature points from a single image of intensity data.

Indexing systems typically require the feature sets to be of some minimum cardinality to perform their function correctly. A previous indexing system for indexing general three-dimensional model points from two-dimensional image data [4] required the point sets to consist of at least four points and each set was represented on a two-dimensional surface in a four-dimensional table. By using a probabilistic method that allows false negatives (matches that are correct, but are not indexed), a system has been created that can perform indexing using point sets with cardinality three and that represents each set in a single bucket in a two-dimensional look-up table.

This work uses the probabilistic peaking effect [1], [2], [3] to discriminate between likely and unlikely matches. The principle of the probabilistic peaking effect is that angles and ratios of distances between points in the model do not vary much when projected onto the image as the viewpoint varies over much of the viewing sphere. This means that the probability density functions of these angles and ratios of distances of projected (image) points have a strong peak at the preprojection (model) value. This effect has been used to build a probabilistic indexing system that can index using point sets with cardinality three.

The ability to index using point sets with cardinality three is important. If there are n image points and m model points, then there are $O(n^k)$ sets of image points and $O(m^k)$ sets of model points with cardinality k. Thus, reducing the set cardinality necessary reduces the number of such sets to consider immensely. In addition, several algorithms (e.g., [5], [11], [13]) use initial matches of three image points to three model points since this is the minimum number necessary to constrain the number of transformations that align the points to a finite set. Current indexing systems that require sets larger than three points cannot generate ideal candidate matches for these algorithms.

While this method has the advantage that smaller image and model groups can be used, it has the disadvantage that we will not index all of the correct model groups. This will be overcome by ensuring that we examine several correct hypotheses.

Let us call the sets of image points that have been grouped together for use in indexing the table *image groups* and the sets of model points that are hypothetically matched to them *model groups*. If each of the points in an image group is the projection of the corresponding point in a model group, then we will say that the two groups are in *actual correspondence*. The premise of this system is that the probabilistic peaking effect is a strong enough indicator of image feature values to eliminate the vast majority of model groups that are not in actual correspondence with a specific image group while indexing the one that is a significant fraction of the time.

II. Indexing for Object Recognition

Indexing systems for machine vision generate an index vector from a set of image features. This vector is used to look up the sets of model features that could have projected to the image features in a multidimensional index table. Ideally, one is able to represent a set of model features by an index vector that remains the same regardless of transformation or projection. Such an index vector is said to be *invariant*.

Once an invariant vector has been found, an index table can be created by discretizing the vector space of invariant parameters. Model features are then stored in the index table at the locations corresponding to their invariant vector. At run-time, the index vector associated with the image features can be used to look up the model features in the index table. Image noise complicates the indexing process considerably.

Invariants have been found for several model representations. For example, Lamdan et al. [12] describe invariants of two-dimensional point sets (with cardinality four or more) undergoing affine transformations and orthographic projection. Forsyth et al. [8] describe projective invariants of two-dimensional algebraic curves (e.g., conics.) Differential invariants of general two-dimensional curves are given by Weiss [17].

It has been proven that no invariants exist for single views of general three-dimensional points sets [3], [4]. Despite this, Clemens and Jacobs [4] have shown that an indexing system for this problem can be built that (in the noiseless case) indexes exactly those groups that could have projected to a specific image group under weak-perspective. This system requires groups to consist of (at least) four points.

Since Huttenlocher and Ullman [11] have shown that any set of three non-colinear model points can be brought into alignment with any three image points by a weak-perspective transformation, it is clear that no conventional indexing system can perform indexing for this case using sets of three points. This property also holds for perspective projections.

III. The Probabilistic Peaking Effect

While it has been proven that there is no affine or projective invariant for general three-dimensional point sets, it has been observed that there is a strong peaking effect in the probability densities of many angles and ratios of lengths in images at the values taken by the features in the model [1], [2], [3]. This means that there is a large range of viewing directions over which these values change in the image by a small amount. This information can be used to discard matches between image groups and model groups that have a small likelihood of being in actual correspondence.

Consider a model group that has been projected onto the image plane. Let p_1, p_2, and p_3 be the points in the model group and q_1, q_2, and q_3 be the corresponding image points. Also, let α be the angle

Manuscript received May 19, 1993; revised Oct. 17, 1994.
C.F. Olson is with the Department of Computer Science, Cornell University, Ithaca, NY 14853 USA; e-mail: clarko@cs.cornell.edu.
IEEECS Log Number P95049.

$\angle p_1p_2p_3$ and β be the angle $\angle q_1q_2q_3$. Define the segment lengths as follows: $a_1 = \left|\overline{p_1p_2}\right|$, $a_2 = \left|\overline{p_2p_3}\right|$, $b_1 = \left|\overline{q_1q_2}\right|$, $b_2 = \left|\overline{q_2q_3}\right|$. See Fig. 1. The features that are used in this system to determine which groups are most likely to match are:

1) The angles formed by the points in the model (α) and in the image (β).
2) The ratios of the lengths of the segments in the model $\left(\frac{a_1}{a_2}\right)$ and in the image $\left(\frac{b_1}{b_2}\right)$.

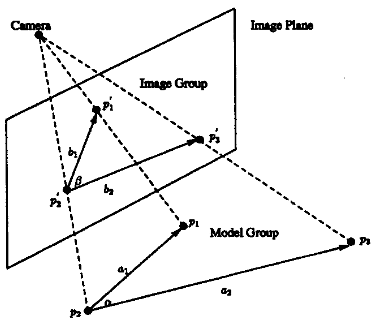


Fig. 1. A model group projected onto the image plane using the perspective projection.

Ben-Arie [1] gives an equation to approximate the probabilistic peaking effect of these features as it varies with $\frac{\beta}{\alpha}$ and $\frac{b_1a_2}{b_2a_1}$. It should be noted that this peaking effect varies not only with these ratios, but also with α (or alternatively with α, β, and $\frac{b_1a_2}{b_2a_1}$). Ben-Arie's approximation of the joint probability density does not model this effect. To better model the probabilistic peaking effect, the probability histograms have been recreated through numerical integration with the additional variable α. As in the experiments performed by Ben-Arie, the viewing sphere was tessellated and the area of each tessellation was added to the bucket corresponding to the image angle β and the ratio of lengths $\frac{b_1a_2}{b_2a_1}$ from the viewing direction at the center of the tessellation.

In addition, previous models for the probabilistic peaking effect have not modeled localization error in the feature detection process. To account for such noise, the probability histograms have been generated with bounded noise ($\varepsilon = 1.0$ and $\varepsilon = 3.0$) added to the image parameters. The bounded noise model specifies that the true location of each image feature is within some distance ε of the measured location. Fig. 2 shows examples of the joint probability histograms for the case with noise ($\varepsilon = 1.0$).

IV. Probabilistic Indexing

The probabilistic peaking effect can be used to create a system that can determine through indexing which model groups are likely to have projected to specific image groups. The first step is to create a look-up table containing the features values (α and $\frac{a_1}{a_2}$) for each model group.

To determine which model groups are the most likely to have pro jected to an image group, we search the probability histograms described in the previous section. This search determines a subset of buckets in the index table that is called a cloud due to its shape. To

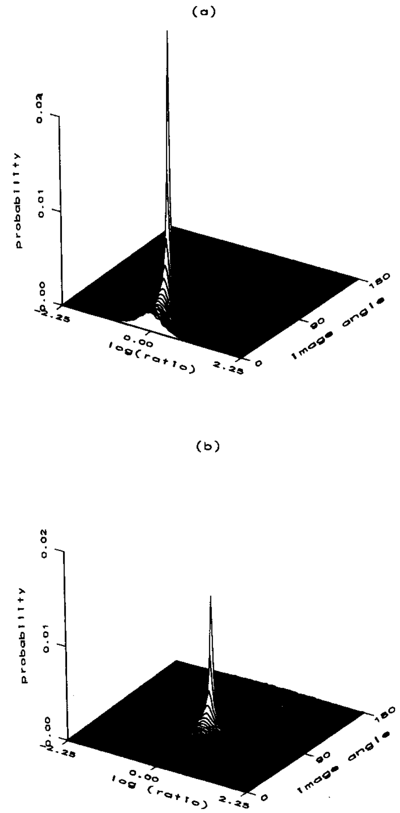


Fig. 2. The joint probability histograms showing the probabilistic peaking effect for selected values of the model angle α with noise ($\varepsilon = 1.0$). The x-axis is the image angle β. The y-axis is the logarithm of the ratio of lengths $\log\frac{b_1a_2}{b_2a_1}$. The z-axis is the probability. (a) $\alpha = 30°$; (b) $\alpha = 90°$.

determine the extent of the cloud we vary the model angle α and the ratio $\frac{b_1a_2}{b_2a_1}$ and determine which cells of the histogram have values greater than some predetermined threshold. We do not need to vary the angle β because this is fixed by the image group. Each histogram bucket then corresponds to one or more index table bucket. (We take the range of $\frac{b_1a_2}{b_2a_1}$ above the threshold for some α and divide it from $\frac{b_1}{b_2}$ to get the range of index table buckets.) We need not worry about noise in the image features when indexing because we have already accounted for it in the probabilistic peaking effect probability histograms. Each model group contained in each bucket of the cloud is considered as a possible match for the current image group.

The performance of probabilistic indexing has been estimated for various probability thresholds through experimentation on random point sets. These experiments transformed the models by a random three-dimensional rotation and projected them onto the image using the perspective projection. Bounded noise ($\varepsilon = 1.0$) was added to coordinates of each image point. Table I shows the fraction of correct matches and incorrect matches that were indexed for various probability thresholds.

If we know the prior probability distribution of image group fea ures, we can use Baye's rule to determine the posterior probability of each match being correct. Let $b = b_i$ be the proposition that the bin the image group falls in is b_i and let h be the proposition that the model group and the image group are in actual correspondence.

TABLE I
FRACTION OF CORRECT AND INCORRECT MATCHES INDEXED FOR VARIOUS PEAKING PARAMETER THRESHOLDS WITH NOISE. T IS THE PROBABILITY THRESHOLD USED TO DETERMINE WHICH MATCHES ARE INDEXED, ρ IS THE FRACTION OF CORRECT MATCHES INDEXED, p IS THE FRACTION OF INCORRECT MATCHES INDEXED.

T	ρ	p
.001	.468	.0655
.002	.335	.0277
.003	.265	.0160
.004	.226	.0109
.005	.193	.0081
.006	.173	.0062
.007	.159	.0051
.008	.140	.0042
.009	.128	.0035
.010	.115	.0029

$$P(h \mid b = b_i) = \frac{P(h)P(b=b_i \mid h)}{P(b=b_i)}$$

$P(b = b_i \mid h)$ is given by the probability histograms of the probabilistic peaking effect. $P(b = b_i)$ is given by the prior probability histogram of image feature values. This histogram has been generated through numerical integration similar to the peaking effect histograms. See [15] for details. I have assumed that the prior probability of each possible match (and thus each possible hypothesis h) is the same, so we can drop the $P(h)$ term without changing the ranking of the hypotheses. Of course, if we had knowledge that model groups were not equally likely to appear in the image, we could use it here.

Probabilistic indexing has been extended to point sets with cardinality greater than three. Details can be found in [15] and a discussion of the merits of using larger groups can be found there or in [16].

V. FAST ALIGNMENT

Probabilistic indexing techniques can be used to improve the performance of many recognition systems. For example, this section will discuss how these techniques can be used to speed up the alignment method [11].

The alignment method is a model-based object recognition technique for recognizing three-dimensional objects from a single view in intensity images. The key to the alignment method is that a unique (up to a reflection) weak-perspective transformation between the model and the image of the model can be found by matching three model points to three image points. To recognize objects, all triples of image points can be hypothetically matched to all triples of model points. For each such match, the transformations that bring them into alignment are determined and each of these transformations is then tested to determine if it is correct.

If there are m model points and n image points, this would require $O(m^4n^3 \log n)$ operations, since the testing step is $O(m \log n)$. It is thus not desirable to examine each possible match between three image points and three model points. If an object is present in the image, it is likely that a substantial number of triples of model points can be detected. In the best case, only one of these triples needs to be found and matched to recognize the object. If all of the matches are examined, then much work is being done that is not necessary. Even if we stop once a adequate match has been found, we can use information about the likelihood of each match being correct to determine which matches are best to examine.

Speedup can be achieved by using probabilistic indexing to determine which matches are most likely to be correct. These matches are tested and the rest are not considered. Error criteria are given in [14] that can be used to further reduce the number matches that need be examined.

A. Speedup

We can use the results from random points sets to estimate the speedup and the probability of a false negative when using probabilistic indexing in conjunction with the alignment method. Let's examine the speedup achieved with two different stopping criteria:

1) All image groups are used for indexing. Each match between an image group and a model group that is indexed is tested and the best match is accepted as correct if it meets some criterion.
2) Image groups are examined in some order. Matches that are indexed are tested. As soon as one of these matches meets some criterion, it is accepted as correct and the remainder of the matches are not examined.

Let's define the speedup as the expected number of transformations that must be tested by the algorithm when not using probabilistic indexing divided by the expected number when using probabilistic indexing. I do not consider the overhead necessary to determine if a match is indexed, since this process is a fast, constant time operation (when amortized over the matches) and the testing step is $O(m \log n)$. Let g be the total number of matches considered, p be the fraction of total matches indexed, γ be the number of correct matches considered, and ρ be the fraction of these matches that are indexed.

With the first stopping criterion, we examine gp matches when using probabilistic indexing and g matches when not using probabilistic indexing, so the speedup is simply $\frac{1}{p}$.

With the second stopping criterion, if we assume that matches are examined in random order, we have a hypergeometric distribution. For large values of g, this can be approximated by the binomial distribution. The expected number of matches that must be tested in this case when not using probabilistic indexing is approximately $\frac{g}{\gamma}$. When using probabilistic indexing, the expected number of matches that must be tested is approximately $\frac{pg}{\rho\gamma}$. The speedup is thus approximately $\frac{\rho}{p}$. This analysis assumes that $\gamma > 0$, that is, a correct match exists. If $\gamma = 0$, as is the case when the model is not present in the image, the speedup is the same as for the first stopping criterion.

TABLE II
SPEEDUPS FOR VARIOUS INDEXING THRESHOLDS: T IS THE INDEXING THRESHOLD, ρ IS THE FRACTION OF CORRECT MATCHES INDEXED, p IS THE FRACTION OF INCORRECT MATCHES INDEXED, $\frac{\rho}{p}$ IS THE SPEEDUP FOR STOPPING CRITERION 2 WHEN $\gamma > 0$, $\frac{1}{p}$ IS THE SPEEDUP FOR STOPPING CRITERION 1 AND STOPPING CRITERION 2 WHEN $\gamma = 0$.

T	ρ	p	$\frac{\rho}{p}$	$\frac{1}{p}$
.002	.335	.0277	12.1	36.1
.005	.193	.0081	23.8	123.5
.008	.140	.0042	33.3	238.1
.010	.115	.0029	39.7	344.8

Table II shows ρ, p, and the expected speedups for some values of the indexing threshold. Large speedups are attained for the first stopping criterion and for the second stopping criterion when $\gamma = 0$. The speedups for the second stopping criterion when $\gamma > 0$ are not as large, but we can still use these techniques to speed up the recognition process by up to a factor of 40. Note that the case $\gamma = 0$ is the common case, since each model in the database must be considered and there are usually few of the models present in the image.

The speedup achieved through the use of probabilistic indexing is a constant factor speedup. The computational complexity of the algorithm remains the same. It should be noted though, that several studies have shown that feature indexing systems index a constant fraction (on average) of the feature groups in the index table in the presence of noise [4], [9], [10]. If each of the indexed groups requires at least $O(1)$ time to process, we cannot achieve more than a constant factor speedup through indexing, despite what some authors have claimed.

If we assume that the probability of a correct match being indexed is independent of whether other correct matches are indexed (this assumption will be discussed below), the probability of a false negative as a result of not indexing any correct matches is $(1-\rho)^{\gamma}$ for both stopping criteria.

TABLE III
PROBABILITY OF A FALSE NEGATIVE FOR VARIOUS VALUES OF ρ (THE FRACTION OF CORRECT MATCHES INDEXED) AND γ' (THE NUMBER OF CORRECT MATCHES).

ρ	$\gamma' = 25$	$\gamma' = 50$	$\gamma' = 100$
.115	4.72×10^{-2}	2.22×10^{-3}	4.95×10^{-6}
.140	2.30×10^{-2}	5.31×10^{-4}	2.82×10^{-7}
.193	4.70×10^{-3}	2.21×10^{-5}	4.87×10^{-10}
.335	3.72×10^{-5}	1.38×10^{-9}	1.91×10^{-18}

If even 10 points from a model are present in an image, there are 120 correct matches (for 20 model points there are 540 correct matches). Not all of these will result in a correct identification of the object, due to noise, but many will. Let γ' be the number of correct matches that result in correct identification of the object. Table III shows the probability of a false negative for the values of ρ from Table II and $\gamma' = 25$, 50, and 100. Even for relatively small values of γ' and ρ, there is a small chance of a false negative. As γ' or ρ increases, the likelihood of a false negative becomes negligible.

Let's now consider the question of whether the probabilities of different correct matches being indexed are independent. Specifically, we want to know if it is possible for some object to be in an orientation from which all model groups appear in unlikely configurations in an image and are thus not indexed. Unlikely configurations occur when the viewing direction is such that the model group is considerably foreshortened. This means that groups of coplanar points will have unlikely configurations from the same viewing directions. Objects that are not nearly flat will not be a problem since they have groups in a wide variety of orientations. The performance of probabilistic indexing will also be good on flat or nearly flat objects that are not considerably foreshortened, since all or almost all of the correct matches will be indexed. When such objects are rotated such that they are considerably foreshortened in the image, they produce angles and distance ratios far from the probability peaks and thus will be difficult to recognize using probabilistic indexing techniques.

B. Real Images

Probabilistic indexing has been tested on real images of several planar and nonplanar objects. Feature points were detected in the images using a fast and precise interest operator [6], [7]. Indexing was then performed using these points. Table IV shows the fraction of correct and incorrect matches that were indexed for several of these images with T = .005. The fractions of correct and incorrect groups indexed in these images are consistent with the fractions determined using synthetic data (r = .193 and p = .0081). The third object in Table IV (called "cross") is a planar object and the fractions are given in decreasing order of the foreshortening of this object. The image with the maximum foreshortening is rotated such that slant of the cross is approximately 60°. Thus, even when there is significant foreshortening, enough correct groups were indexed by probabilistic indexing to recognize this object.

TABLE IV
INDEXING PERFORMANCE ON REAL IMAGES FOR $T = .005$: ρ IS THE FRACTION OF CORRECT MATCHES INDEXED, p IS THE FRACTION OF INCORRECT MATCHES INDEXED.

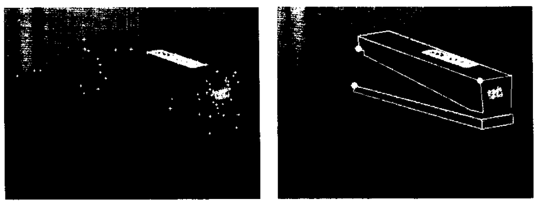
Image	ρ	p
stapler1	.250	.0086
stapler2	.214	.0087
stapler3	.257	.0101
stapler4	.302	.0086
stapler5	.152	.0094
stapler6	.179	.0084
disk1	.247	.0082
disk2	.216	.0075
disk3	.226	.0076
disk4	.456	.0080
cross1	.154	.0085
cross2	.325	.0086
cross3	.781	.0085

Figs. 3 and 4 show examples of recognition being performed using these techniques. Fig. 3(a) shows the corners found in an image of a stapler. While many of the modeled stapler corners were found, several were not. In addition, many unmodeled features were found in this image. The circles in Fig. 3(b) show the points of one of the correct groups that was indexed. The outline of the stapler that is drawn in is the stapler model after being transformed such that the correct model points were aligned with the three image points by a weak-perspective transformation as given by Huttenlocher and Ullman [11]. Of course, many other correct and incorrect groups were indexed. The verification techniques described by Huttenlocher and Ullman can be used to determine which of the matches are correct. For this image, a speedup of 29.1 was achieved in recognizing the stapler using the second stopping criterion. A speedup of 116.3 was achieved in determining that no other objects were present in the image.

Fig. 4 shows a more complicated scene with two objects in it. The corners that were found in this image are shown in Fig. 4(a). Fig. 4(b) shows a correct group that was indexed for each model and the model poses that were determined for them. While the stapler pose is quite good, the pose of the disk is mediocre due to the perspective effects in the scene. The disk is not far enough from the camera for these effects to be negligible and the pose determined under weak-perspective is not ideal. Note that this is not a problem with probabilistic indexing. For this object, 21.6% of the correct groups were indexed, despite the perspective distortion. To alleviate the problem, we could use probabilistic indexing with an algorithm that models the perspective projection (e.g., [5], [13]). For this image, a speedup of 24.6 was achieved in recognizing the stapler, a speedup of 30.1 was achieved in recognizing the disk, and a speedup of 117.6 was achieved in determining that no other objects were present in the image.

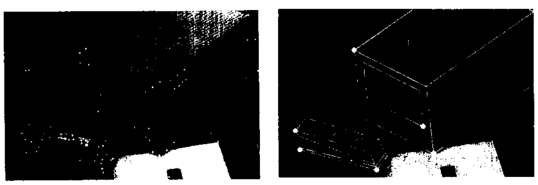
VI. SUMMARY

This correspondence has described an indexing system for use in recognizing three-dimensional objects in single two-dimensional images. The probabilistic peaking effect has been shown to be effective for use in indexing model groups undergoing general rigid transformations in three-dimensions from image points generated using the perspective projection. Its use has allowed us to reduce the cardinality of the sets of image and model points necessary in an indexing system, while retaining the indexing speedup. The disadvantage to this system is that not all correct matches between image groups and model groups are indexed. Since a far higher fraction of correct matches are indexed than of incorrect matches, probabilistic indexing can be used to help discriminate between correct and incorrect hypotheses. These techniques have been applied to the alignment method and found to speed up the recognition process by a considerable amount.



(a) (b)

Fig. 3. Recognition of a stapler: (a) The corners found in the image; (b) A correctly indexed group and the corresponding model pose.



(a) (b)

Fig. 4. Recognition in a more complicated scene: (a) The corners found in the image; (b) Correctly indexed groups and the corresponding model poses.

AKNOWLEDGMENTS

This work has been supported by a National Science Foundation Graduate Fellowship to the author, NSF PYI Grant IRI-8957274 to Jitendra Malik, and NSF Materials Handling Grant IRI-9114446.

REFERENCES

[1] J. Ben-Arie, "The probabilistic peaking effect of viewed angles and distances with application to 3D object recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 12, pp. 760-744, Aug. 1990.

[2] T.O. Binford, T.S. Levitt, and W.B. Mann, "Bayesian inference in model-based machine vision," *Uncertainty in Artificial Intelligence 3*, L.N. Kanal, T.S. Levitt, and J.F. Lemmer, eds., pp. 73-95. North Holland: Elsevier Science Publishers B.V., 1989.

[3] J.B. Burns, R.S. Weiss, and E.M. Riseman, "View variation of point-set and line segment features," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 15, pp. 51-68, Jan. 1993.

[4] D.T. Clemens and D.W. Jabobs, "Space and time bounds on indexing 3D models from 2D images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 13, pp. 1,007-1,017, Oct. 1991.

[5] M. A. Fischler and R.C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Comm. ACM*, vol. 24, pp. 381-396, June 1981.

[6] W. Förstner, "Image matching," chapt. 16 of *Computer and Robot Vision*, vol. II, by R. Haralick and L. Shapiro, Addison-Wesley, 1993.

[7] W. Förstner and E. Gülch, "A fast operator for detection and precise locations of distinct points, corners, and centers of circular features," *Proc. Intercommision Conf. Fast Processing of Photogrammetic Data*, pp. 281-305, 1987.

[8] D. Forsyth, J.L. Mundy, A. Zisserman, C. Coelho, A. Heller, and C. Rothwell, "Invariant descriptors for 3D object recognition and pose," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 13, pp. 971-991, Oct. 1991.

[9] W.E.L. Grimson and D.P. Huttenlocher, "On the sensitivity of geometric hashing," *Proc. Int'l Conf. Computer Vision*, pp. 334-338, 1990.

[10] W.E.L. Grimson, D.P. Huttenlocher, and D.W. Jacobs, "A study of affine matching with bounded sensor error," *Proc. European Conf. Computer Vision*, pp. 291-306, 1992.

[11] D.P. Huttenlocher and S. Ullman, "Recognizing solid objects by alignment with an image," *Int'l J. Computer Vision*, vol. 5, no. 2, pp. 195-212, 1990.

[12] Y. Lamdan, J.T. Schwartz, and H.J. Wolfson, "Object recogniton by affine invariant matching," *Proc IEEE Conf. Computer Vision and Pattern Recognition*, pp. 335-344, 1988.

[13] D.G. Lowe, "Three-dimensional object recognition from single 2D images," *Artificial Intelligence*, vol. 31, pp. 355-395, 1987.

[14] C.F. Olson, "Fast alignment using probabilistic indexing," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 387-392, 1993.

[15] C.F. Olson, "Probabilistic indexing: Recognizing 3D objects from 2D images using the probabilistic peaking effect," Tech. Report UCB//CSD-93-733, Univ. of California at Berkeley, May 1993.

[16] C.F. Olson, "Probabilistic indexing: A new method of indexing 3D model data from 2D image data," *Proc. Second CAD-Based Vision Workshop*, pp. 2-8, 1994.

[17] I. Weiss, "Noise resistant projective and affine invariants," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 115-121, 1992.

Color Constant Color Indexing

Brian V. Funt and Graham D. Finlayson

***Abstract*—Objects can be recognized on the basis of their color alone by *color indexing*, a technique developed by Swain and Ballard [15] which involves matching color-space histograms. Color indexing fails, however, when the incident illumination varies either spatially or spectrally. Although this limitation might be overcome by preprocessing with a color constancy algorithm, we instead propose histogramming color ratios. Since the ratios of color RGB triples from neighboring locations are relatively insensitive to changes in the incident illumination, this circumvents the need for color constancy preprocessing. Results of tests with the new color-constant-color-indexing algorithm on synthetic and real images show that it works very well even when the illumination varies spatially in its intensity and color.**

***Index Terms*—Color indexing, color constancy, retinex, object recognition.**

I. INTRODUCTION

Swain and Ballard [15] developed a very clever, simple scheme that identifies objects entirely on the basis of color. Their method, which they call *color indexing*, radically departs from traditional object recognition strategies based on geometric properties. Color indexing turns out to be remarkably robust in that variations such as a change in orientation, a shift in viewing position, a change in the scene background, partial occlusion, or even a radical change in shape (e.g., a shirt tossed onto a chair two different ways), degrade recognition only slightly.

On the other hand, Swain's algorithm is very sensitive to the lighting. Simple changes in the illumination's intensity—let alone its color—radically alter the algorithm's results. Clearly, one solution to

Manuscript received Feb. 19, 1992; revised Aug. 27, 1993.

The authors are with the School of Computing Science, Simon Fraser University, Vancouver, B.C., Canada V5A 1S6; (604) 291-3126; e-mail funt@cs.sfu.ca.

IEEECS Log Number P95047.