TABLE 3
RESULTS OF 4,000 ESTIMATES
OF $\beta$ IN EXPERIMENTS 1 AND 2

<table>
<thead>
<tr>
<th>Noise Level</th>
<th>0.000</th>
<th>0.005</th>
<th>0.010</th>
<th>0.020</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\beta}$ (Experiment 1)</td>
<td>13.305±4.150</td>
<td>12.304±2.152</td>
<td>13.320±4.259</td>
<td>13.357±8.647</td>
</tr>
<tr>
<td>$\hat{\beta}$ (Experiment 2)</td>
<td>13.306±4.150</td>
<td>12.306±2.154</td>
<td>13.320±4.258</td>
<td>13.327±8.733</td>
</tr>
</tbody>
</table>

8 CONCLUSION
In this paper, a new method for finding motion from three weak perspective images has been presented. This method relies on rotating a second image in order to simplify the rotation matrix relating it to an initial image. A function has also been provided that allows one to easily achieve a correct image rotation using point correspondence between this image pair. Also provided, was a linear algorithm to solve for the unknown elements of both rotation matrices once this second image has been correctly rotated and a third image has been introduced.

REFERENCES

An Integration Scheme for Image Segmentation and Labeling Based on Markov Random Field Model
II Y. Kim and Hyun S. Yang

Abstract—This paper presents a unified approach for the image understanding problem based on the MRF models. In the proposed scheme, the image segmentation and interpretation processes cooperate in the simultaneous optimization process so that the erroneous segmentation and misinterpretation can be compensately recovered by continuous estimation of the unified energy function.

Index Terms—Region clustering, region labeling, Markov random field, energy function, optimization.

1 INTRODUCTION
In general, image analysis recognizes, locates the position and orientation of, and provides a sufficiently detailed symbolic description or recognition of imaged objects deemed to be of interest in the three-dimensional environment. Among those various tasks involved in image understanding, we focus on region analysis which tries to segment an input image into meaningful regions and assign object labels to each of them.

In general, the initial set of segmented regions is not complete, that is, a region of the set does not always exactly match to a real object surface. Thus, recognizing those regions without any modification on the initial set is quite errorful. Contrarily, the exact separation of the given scene into true object surfaces is not usually available if we do not use the task specific knowledge about the scene at all. Thus to get more optimal set of segmented regions and interpretation results, the constraints about the image segmentation and interpretation should be applied to the scene understanding process cooperatively.

Although the previous works [1], [2], [3] and others have been successfully demonstrated to various degrees thus far in incorporating the interpretation process into the segmentation process, they still follow the similar fashion with the region growing technique by propagating the region analysis from the regions with established judgement to the nearby regions.

The integration of various constraints for solving the given problems in formal scheme has recently been studied with focuses on the Markov Random Field model [4], [5]. Although most of the works related with the MRF model dealt with the pixel-level image structures as the sites to be labeled, the natural extension to the more abstract level of image structures such as regions can be achieved by defining an MRF model on and between regions. Modesto and Zhang [6] first applied the MRF-based region labeling scheme to the image interpretation problem.

* I.Y. Kim is with the Micro Devices Business, Samsung Electronics Co. Ltd., 82-3 Dodang-Dong, Wommi-Ku, Bucheon, Kyunggi-Do, Korea 421-130. E-mail: iykim@brain.koist.ac.kr.
* H.S. Yang is with the Department of Computer Science, KAIST, 373-1 Kusung-Dong, Yusung-Ku, Taejon, Korea 305-701. E-mail: hyang@cs.kaist.ac.kr.

Manuscript received Aug. 25, 1994; revised August 1995. Recommended for acceptance by R.L. Kashyap.
For information on obtaining reprints of this article, please send e-mail to: transactions@computer.org, and reference IEEECS Log Number P95157.
Fig. 1. An example of mislabeling results due to missegmentation: (a) a scene; (b) boundaries of initial set of segmented regions; (c) boundaries of enhanced set of segmented regions; and (d) labeling results.

In the work proposed here we also consider the regions as the basic elements to be labeled by the MRF-based site labeling scheme. But in this work we put an emphasis upon the interaction between distinct levels of labeling process, the segmentation and interpretation via an intermediate level of data structures called region clusters. That is, we initially segment a given image into a preliminary set of segmented regions. The segmentation problem is then defined to group those regions into an optimal set of region clusters with respect to the given constraints for segmentation while the interpretation process finds an optimal set of interpretation labels for the given set of region clusters.

The rest of the paper is organized as follows. In Section 2, we present an overview of the proposed integrated image understanding scheme. Section 3 deals with the MRF-based problem formulation including the specific design of the unified energy function for encoding the constraints of the image segmentation and interpretation problem. In Section 4, we present the experimental results on real outdoor scenes together with the comparisons between sequential and integrated scheme. Finally, we conclude in Section 5.

2 An Overview of the Proposed Scheme

Before explaining the proposed integrated scheme for region segmentation and interpretation, we need to consider the reason that we take the integrated approach.

2.1 Motivations

Fig. 1a shows an image describing outdoor scene. As you see in the image, the road is very bright at its upper-right part and the darker its brightness is the farther it is from the upper-right part due to the lighting condition. This style of changes of brightness causes a single object surface to be divided into several disjoint regions in traditional region-based segmentation techniques. Fig. 1b shows it when we segment the given image by one of the region-growing segmentation techniques [7].

An effort to enhance those incomplete sets of segmented regions based on MRF model has already been done in our early work [8]. However, in that work, we considered the optimality only in the respect of image segmentation. Although the segmentation results could be achieved to satisfy the given constraints, it still does not behave as a complete segmentation scheme. This can be seen in the Fig. 1c, where the road remains separated into two regions which could not be merged since there exists strong disparities in spectral features between two road regions. And thus the spatial relation between the road and roadlines has been misrepresented to the interpretation process and those regions are labeled as different labels (Fig. 1d) when we applied our MRF-based image interpretation scheme. In the Fig. 1d, the patterns are designed to distinguish interpretation labels and equally defined as in [9].

2.2 The Global Scheme

The proposed integrated scheme is shown in Fig. 2. In the figure, the left-half and the right-half of the figure corresponds to the image segmentation and interpretation process as in [8], [9], respectively.

Fig. 2. An integrated scheme for image segmentation and interpretation process.

As before, we start with the initial set of segmented regions by applying the ISM (Initial Segmentation Module) which is implemented as one of the traditional region-based image segmentation technique. In the integrated scheme, the interpretation process is then applied to the temporal set of region clusters which are formed from the current set of region labels assigned to the initial set of segmented regions by the segmentation process. The optimality of the current sets of region labels for segmentation and interpretation is then jointly estimated through a unified energy function.

But at the initial stage of the optimization process, the current set of region clusters may be far from the ideal segmentation state, which indisputably misleads the interpretation process. And also the misinterpretation labels for the current set of region clusters also prevent the segmentation process from proceeding to the more optimal segmentation state via the unified energy function. Thus we should control the weights between two processes as the optimization process proceeds. That is, at the initial stage of the optimization process, we give more attention to the segmentation process so that it almost build up the optimal segmented state. And then, as the optimality of the interpretation labels increases, we raise the importance of the interpretation process by degrees. The weights for both processes are determined by adaptive weighting functions which monitor the optimality of the current sets of region segmentation.
and interpretation labels. In the Fig. 2, those weighting functions are denoted as $I_r, I_p$ and resistance symbols.

3 Problem Formulation

Now before formulating the given image segmentation and interpretation problem within an integrated scheme, we should define an MRF on the sites to be labeled. Since we basically follow the scheme by Modenstino and Zhang [6] in defining the MRF on and between regions, we will not explain the underlying principles of MRF definitions but directly present the key descriptions of our MRF-based integration scheme such as sites to be labeled, set of labels and energy functions to be minimized.

3.1 Configurations

Let $S_t$ be a set of sites or regions for the segmentation process, $X_t$ a set of random variables for denoting the labels assigned to the sites in $S_t$, $A_t$ a set of labels such that each site in $S_t$ can have, and $o_t$ be one of instances from configuration space, $\Omega_t$. And also let $S_t$ be a set of sites in interpretation process, $X_t$ a set of random variables associated with each element in $S_t$, and $\Lambda_t$ be a set of object labels that each site in $S_t$ can have. Since the set of region clusters is continuously reconstructed by temporally merging neighboring regions with the same labels in $\Lambda_t$, $S_t$ can be denoted as $S^t$, where $\omega^t$ corresponds to a labeling instance generated by the segmentation process for the sites in $S_t$ at time $t$. The problem configuration can then be summarized as:

$$S_t = \{R_{1t}, R_{2t}, ..., R_{nt}\}$$

$$X_t = \{X_1, X_2, ..., X_N\}$$

$$\Lambda_t = \{0, 1, 2, ..., N\}$$

$$S^t = \{P^t_1, P^t_2, ..., P^t_N\}$$

$$X_t = \{Y_1, Y_2, ..., Y_N\}$$

$$\Lambda_t = \{\text{Sky, Foliage, Road, ... Tree}\}$$

where $N$ is the number of initially segmented regions, each $P^t_i$ a region cluster constructed from the labeling configuration $\omega^t$ and $N_t$ is the number of region clusters at time $t$.

3.2 Unified Energy Function

Before defining the energy function, let's introduce some notations and detailed descriptions of which can be found in [8], [9].

$F_{ir}$ a feature vector representing the unary features on the region $R_i$.

$E^t_{ir}$ the $t$th feature measurement on the region $R_i$.

$G_t$ a set of neighboring regions of $R_i$.

$\Phi_t$ a feature vector representing the unary features on the region cluster $P_i$.

$\Psi_t$ a feature vector representing the binary features measured between region clusters $P_i$ and $P_j$.

$\tau_i$ index of the region cluster that contains the region $R_i$.

Finally, if we let $\hat{\omega}^t$ be a labeling configuration achieved through optimization for the current set of sites $S^t$ in interpretation process, the general form of the unified energy function can then be defined as:

$$U(\omega^t, \hat{\omega}^t) = \sum_{R \in S^t} (J_1 \times E^t_{IR} + J_2 \times E^t_{IP})$$

where $J_1$ and $J_2$ are the adaptive weight functions, $E^t_{IR}$ and $E^t_{IP}$ are the energy terms for segmentation and interpretation, respectively. The $E^t_{IR}$ is specified as:

$$E^t_{IR}(\omega^t) = \sum_{K ightarrow \omega^t} \left( \eta_k \frac{\text{max}(F^t_{IR} - F^t_{IP})}{\text{min}_{\omega^t}(F^t_{IR} - F^t_{IP})} \right) + \alpha(1 - \eta_k)$$

where $\alpha$ is a constant.

The underlying assumption of the above energy function is that in the set of optimally segmented regions, a region should be uniform in the features used and between distinct separated regions should exist large discontinuities in those features. Thus, in the above equation, the first term on the inside of the summation constrains the given set of region clusters to minimize the Within-Class variance while the last term constrains it to maximize the Between-Class variance. The parameter $\alpha$ controls the weight between those two terms.

Now $E^t_{IR}$ is defined to be:

$$E^t_{IR}(\omega^t) = V_t(\omega^t, \Phi_t) + \sum_{K \rightarrow \omega^t} V_{IK}(\omega^t, \Psi_t)$$

where $\lambda, \lambda'$ and $\lambda''$ are all in $\Lambda_t$.

In the above equation, $V_t$ represents the optimality of the label $\lambda$ as a possible label for the region cluster $P_t$, in the respect of unary features while $V_{IK}$ represents the optimality of the labels $\lambda'$ and $\lambda''$ in the respect of binary features between the region clusters $P_t$ and $P'_t$. And those optimality are determined by a set of neural networks which have been learned from examples [9].

The weight functions, $J_1$ and $J_2$ vary in the course of the optimization and are automatically determined by the following equations:

$$J_1 = 1 + \delta \frac{\sum_{K \in S^t_{IR}} E^t_{IR}(\omega^t)}{n(C_1) + n(C_2)}$$

$$J_2 = 1 + (1 - \delta) \frac{\sum_{K \in S^t_{IR}} E^t_{IP}(\omega^t)}{n(C_1) + n(C_2)}$$

where $\delta$ is a constant for weighting two terms in the unified energy function and $n(C_1)$ and $n(C_2)$ is the cardinalities of the sets of cliques which consist of a single region constrained by $V_t$ and a pair of regions constrained by $V_{IP}$ respectively. That is, the fractional part of the equations means the average value of the clique functions and thus the $J_2$ increases as the current optimality of the labeling configuration, $\hat{\omega}^t$, increases.

3.3 The Optimization Process

Now we find a labeling configuration that minimizes the given unified energy function using the simulated annealing procedure [10].
Stage 1: Initialization
1.1 Select an initial temperature $T$
1.2 Set $i$, and $j$, to 2 and 1, respectively
1.3 Assign initial labels to $X_i$ in $X_j$ such that $X_i \in \Lambda_i \forall i$

Stage 2: Annealing
2.1 Select a variable $X_i$ from $X_j$, change its label to any one of $\Lambda_i$
2.2 Reconstruct the $S_i$ according to the newly generated configuration, $\omega_i$
2.3 Estimate $\beta_i$ for the given $S_i$, using simulated annealing
2.4 Calculate the amount of change in the unified energy function, $\Delta U$
2.5 If $\Delta U < 0$ accept the current $\omega_i$
    otherwise, accept it if $e^{\frac{\Delta U}{T}} > \xi$
    where $\xi$ is a randomly generated number over $[0, 1]$
2.6 Repeat 2.1 ~ 2.5 predetermined number of times
2.7 Replace $T$ by $f(T)$ where $f$ is a monotonically decreasing function
2.8 Reestimate $j$, and $i$, based on the current $\omega_i$ and $\omega_j$, respectively

Stage 3: Repeat Stage 2 until energy becomes stabilized

The final labeling configurations $\omega_i$ and $\omega_j$ at the exit of the annealing procedure are taken as the optimal segmentation and interpretation labels, respectively.

4 Evaluations on Real Scenes

Now, we exploit the designed region segmentation and labeling scheme to get the optimal set of segmented regions and labeling results for the given preliminary set of segmented regions.

First the result of applying the integrated scheme to the set of regions in Fig. 1b is shown in Fig. 3. As you see in the Fig. 3, two separated regions in Fig. 1b corresponding to the road are merged into one to be labeled as a "road" and the region at the center of it is labeled as a "centerline" because with respect to the defined unified energy function it is more optimal rather than to recognize them as in the Fig. 1d. In Fig. 4, we show a partial set of our experimental results.

Fig. 3. The resulting set of (a) segmented regions and (b) Interpretation labels, when we applied the integrated scheme to the road scene.

Now, we turn our attention to the behavior of the adaptive weight functions in (2). Fig. 5 shows the average labeling accuracies in experimentations for the sample set of images. In the figure, the labeling accuracies are measured at each pair of values of $E(\omega_i)$ and $E(\omega_j)$ with $j$, and $i$, not multiplied while the annealing proceeds.

Also in the graph, the point at which three axes meet together corresponds to the global minimum of the unified energy function with weight functions excluded. Thus, if we perform the optimization, we reach that point in most cases and get rather high labeling accuracies. But, the highest average labeling accuracy has been achieved not at that point but at the one denoted by the arrow, at which the energy value for the interpretation is minimum (0.0) while for the segmentation is not (0.3). This means that we should sacrifice the optimality of the segmentation by somewhat degrees to get the most accurate labeling result which is the ultimate goal of the image understanding system.

Fig. 4. Experimental results on outdoor scene images: scenes (first column), boundaries of initial set of segmented regions (second column), boundaries of final set of region clusters at the exit of integrated scheme (third column), and final set of interpretation labels (fourth column).
5 Conclusions

In this paper, we have proposed an image segmentation and interpretation scheme based on the Markov Random Field model as a unified and systematic approach for both problem domains. The proposed scheme basically takes a unified formulation for both problem domains. That is, two major processes in image understanding task, segmentation and interpretation, can be integrated into a single stage of process by incorporating constraints for both problem domains into a unified energy function via corresponding clique functions. In this scheme, the image segmentation and interpretation processes can cooperate in the simultaneous optimization process such that the erroneous segmentation and misinterpretation due to incomplete knowledge about each problem domain can be compensately recovered by continuous estimation of the single unified energy function.

The proposed scheme has been exploited to segment and interpret the images from natural outdoor scenes and shown robustness against the incomplete a priori knowledge about the image segmentation and interpretation problems.

References