A Unified Probabilistic Graphical Model based Approach for the Robust Decoding of Color Structured Light Pattern

Chao Yang[1], Fang Liu[2] and Zhan Song[1]

[1]*Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, The Chinese University of Hong Kong Shenzhen, Guangdong Province, China*

[2]*Department of Aviation Ammunition, China Air Force Logistics College, Xuzhou, Jiangsu Province, China*

{chao.yang, zhan.song} siat.ac.cn, ftshlong2005@126.com

Abstract – **Color coding is an important research topic in spatial encoded structured light sensing (SLS). In this study, we propose a novel graphical model based approach for the color pattern decoding task. For efficient color labeling, the color pattern is firstly decomposed into separate binary pattern images. With the labeled pattern elements, a unified probabilistic graphical framework is constructed to represent the pseudorandom pattern as a clique tree structure. The model contains two parts: the Conditional Random Field (CRF) is used to represent the dependences between these local decisions, and the Bayesian network (BN) is applied for the representation of background colors effect. A colorful target is experimented to demonstrate its feasibility. And the 3D reconstructed models based on the decoding results are also provided to show its robustness.**

Index Terms - Probabilistic graphical model, structured light sensing, pseudorandom array, pattern decoding.

I. Introduction

Spatially encoded structured light sensing has been an important means for 3D reconstruction from single image. By encoding the pattern features with its adjacent pattern elements, and analyzing the captured image, the pattern features with unique codeword can be extracted and evaluated. This procedure also named as the pattern decoding. For most spatially encoded SLS, the color elements are usually adopted for the coding strategy. And thus, their decoding tasks are greatly dependent and affected by the target surface color or texture. How to identify the projected color elements from the polluted images is still an open issue in this research topic.

A comprehensive survey of spatial codification strategies is available from [1]. For the coding schemes, geometric primitives, De-Bruijn sequence, pseudorandom arrays, and M-arrays [2-3] have provided spatially encoded SLS with sufficient features to decode the pattern. In our previous works, we also demonstrated that via the pseudorandom coding scheme, by the use of rhombic color pattern, both the feature position and its orientation can be determined [4]. In [5], a real-time decoding algorithm is proposed, which used a region adjacency graph to generate the data points per frame. This spatial decoding method further improves the robustness of structured pattern against textured or low contrast illumination. Color features have been widely explored using different clustering algorithms including agglomerative hierarchical clustering algorithms (AHCA) and elliptical K-mean [6]. Due to the inherent characteristic of K-mean, good initial points are demanded. Different norm-like color spaces have been used to achieve more separated clustering centroids for high accuracy [7-8]. Using adaptive unsupervised clustering algorithm, non-parametric unsupervised algorithm further utilized this scheme to adapt color structured light to more practical experimental environment [9]. However, such unsupervised algorithms demand a diligent parameter tuning and careful selection of the model. And that makes them lack of practices for real applications especially to the targets that have similar color components with the specific designed structured light pattern.

In this work, the previous proposed pseudorandom color pattern in [4] is adopted and investigated for its robust decoding. A unified probabilistic graphical model is presented for its decoding task. In the algorithm, the CRF is utilized to capture dependences between pattern elements, and the BN method is used to capture the causal effect from object color. By combining the two approaches, a local probabilistic model can be obtained. The discriminative model can simplify the decoding problem into a linear sliding window process. Experiment with a colorful target is used to demonstrate its decoding efficiency and the 3D reconstruction result.

II. Unified Graphical Model for Color Pattern Decoding

In this section, we first describe how we obtain the labelled pattern elements. Secondly, details of the unified graphical model construction are introduced. We show how the EM algorithm is used to integrate BN part with CRF by consider scene color as a latent variable so as to finally achieve one-shot 3D reconstruction. Finally, mathematical details of parameter learning process are provided.

A. Automatic Labelling of the Pattern Features

This work was supported in part by the National Natural Science Foundation of China (NSFC, grant no. 61002040, 61375041), the Introduced Innovative R&D Team of Guangdong Province-Robot and Intelligent Information Technology R&D Team(grant no. 201001D0104648280), and the Shenzhen Key Lab for Computer Vision and Pattern Recognition (grant no. CXB201104220032A).

978-1-4799-4808-6 /14/$31.00 ©2014 IEEE

We first discuss the particular difficulties in getting labelled data in SLS. Then we show how a projection scheme, which is similar to temporal codification, liberates us from hand labelling the data. The adopted pattern contains 65×63 pattern elements. It's a hard work to execute the label process manually one by one. To simplify this procedure, the original color pattern is decomposed into four different binary pattern images, and projected to the scene sequentially. The following work is to segment the binary images and identify their corresponding colors.

Over-segmentation of an image into super-pixel regions is a standard preprocessing step to construct a CRF model for image segmentation [10]. For targets like a horse, cow, sky, ground truth labelling is relatively easy to accomplish. In the application of color spatial SLS, the targets is covered with thousands of color blocks, that makes the manually labelling work in practical. In this work, we first the Normalized Cut method [11] to segment the image into super-pixel regions. And then the K-mean clustering algorithm is applied to get finer segmentation result. As shown in Fig. 1, the labeling result of pattern elements in a colorful object is prone to human errors and the background colors. To get an accurate and efficient labeling result, we decomposed the color pattern into individual components. It just like how different levels of codewords are projected in temporal codification and instead projecting elements that represent factors we want to model. For similarity, let assume our object's color is white and we want to label green element in it. We only projected white light in the same location as green element as shown in Fig. 2. With the correspondent locations, we can locate those real green elements in white color object. Using similar approach, we can capture labelled data that represent how blue and red pattern elements are interacted in green color object.

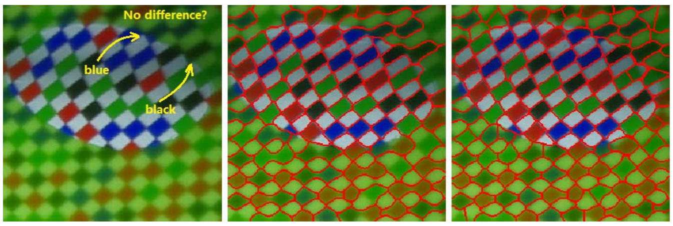

Fig. 1. A colorful surface with color projection (left); Over-segmentation of super-pixel regions by Normalized cut method (middle); Segmentation result by the K-mean clustering based on the NC results (right).

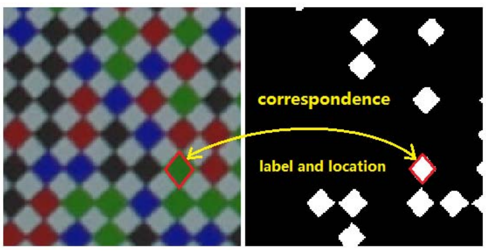

Fig. 2. The colorful pattern is decomposed into 4 separate binary pattern images. Thus, the labeling process can be efficiently solved.

B. *Unified Graphical Model for Color Pattern Description*

The structure of our unified model is as shown in Fig. 3. It's a joint probability distribution of the form:

$$P(C,S) = P(S)P(C_1, C_2, \dots, C_k \mid S, D_i; \theta) \tag{1}$$

It contains two components: CRF part and BN part. Two parts are connected through node C_i , which are the random variables that represent pattern elements' color.

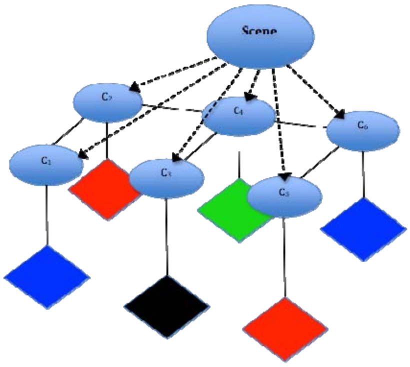

Fig. 3. Diagram of the proposed clique tree structure that used to describe the model of a color pseudorandom structured light pattern.

The proposed unified graphical model is local and with only one unknown S, i.e. the scene color which is association with the BN part latent variable. The CRF part is a common conditional random field over pattern element label C_i given the observation of the image D_i and further condition on scene color S. We use EM algorithm to integrate BN with CRF. In the E-step, it is to assign each pattern element with a soft assignment of pattern element label C. We can infer this conditional probability for each pattern element using the current CRF model parameters. It can be decomposed from the calibrated tree structure (as shown by Fig. 3) of CRF as:

$$P(C \mid S) = P(C_1, C_2, \dots, C_k \mid S, D_i; \theta) \tag{2}$$

Note that k is number of color for the pattern element and this is conditional probability that should be computed for every value of scene color. In the M-step, we estimate our parameters in CRF given the soft assignments of each label from the E-step. Thus, for each of the scene color component, we fit our CRF parameters with each observed pattern element weighted by its conditional probability from Equation (2). It's normal CRF parameters estimation with weighted observed image D_i. In addition to the parameters in CRF, we also need to estimate the scene color distribution $P(S)$ using the expected sufficient statistics obtained in the E-step.

A clique tree structure is constructed to cover all pattern elements. That means we do not consider all the dependences in the codewords. It is just the computational concern. If we put all the dependent factors in the CRF, it will become a graph which results in oscillation and approximation of parameters estimation. With the defined tree structure, it means the Family Preservation and Running Intersection Property [12] hold. So we only need two belief propagations sweep to get a calibration of the distribution before inference about the pattern element labels. Finally, we can use principled algorithms such as the sum-product and the max-product algorithm to perform consistent inference over certain pattern element color or few in calibrated tree. That makes the whole decoding procedure more simplified and efficiently.

C. Model Parameter Estimation

In the E-step, we compute the expected sufficient statistics of each observed image using the conditional probability inferred by the current model parameters. In the M-step, they are aggregated to re-estimate the prior $P(S)$ as:

$$P(S=s)=\frac{M[s]}{\sum_{k=1}^{K} M[k]} \tag{3}$$

Here, $M[s]$ is the aggregated statistics for one possible scene color. In the following section, wherever there is a probability likes $P(C\mid x;\theta)$, it should be $P(C\mid x,s;\theta)$ in which s is the value of latent variable, and x is the weighted observed image D_i in Equation (1).

We build our CRF with log-linear features. A feature can be expressed as a function $f_i(x_i): Val(x_i)\to\mathbb{R}$, where x_i is a set of variables in the scope of the ith feature such as the label of the pattern element and the corresponding observation of the image. Each feature has an associated weight θ_i. Then the CRF distribution can be defined as:

$$P(C\mid x:\theta)=\frac{1}{Z_X(\theta)}\exp\left\{\sum_{i=1}^{k}\theta_i f_i(x_i)\right\} \tag{4}$$

The term $Z_X(\theta)$ is the partition function:

$$Z_X(\theta)\equiv\sum_{C}\exp\left\{\sum_{i=1}^{k}\theta_i f_i(x_i)\right\} \tag{5}$$

We have three types of features:

1) $f_{i,c}^{C}(C_i)$: it operates on single pattern element (an indicator for $C_i=c$);
2) $f_{i,j}^{I}(C_i,x_j)$: it operates on a single pattern element and an image pixel is associated with the label. Here x_j is a statistical measure of the image like mean values of *R*, *G*, *B* channels. They are used to encode the individual probability that $C_i=c$ given x_j;
3) $f_{i,j,c,d}^{P}(C_i,C_j)$: it operates on a pair of adjacent pattern elements (an indicator for $C_i=c, C_j=d$).

Our cost function associated with Equation (1) is negative log-likelihood plus a L_2 regularization penalty on the parameters to prevent overfitting problem as:

$$nll(x,C,\theta)\equiv\log\left(Z_X(\theta)\right)-\sum_{i=1}^{k}\theta_i f_i(C,x)+\frac{\lambda}{2}\sum_{i=1}^{k}\theta_i^2 \tag{6}$$

The traditional BFGS quasi-Newton method is applied to minimize this term for parameters estimation. And then, the CRF model can be used to infer the pattern elements labels in local codeword. Each node *i* is assigned to a label that maximizes its marginal posteriori probability as:

$$c_i=\arg\max_{c_i\in\{r,g,b,black\}} P(C_i\mid x;\theta) \tag{7}$$

The marginal probability $P(c_i\mid x;\theta)$ is calculated by the sum-product belief propagation [12]. In the pattern decoding, the final element labels can be computed using Equation (7) by further marginalizing out the latent variable S.

III. EXPERIMENTS RESULTS

The experimental setup is consisting of one DLP projector and one digital camera with resolution of 1600×1200 pixels. The pseudorandom pattern constructed in our previous work [4] is adopted for the projection. The pattern contains 4 different colors with rhombic shape, and the pattern size is 65×63 with a unique window size of 2×3. The structured light system is calibrated ready by the method in [13].

Fig. 4, complex color object projected with structured light pattern (left) and before that (right). We construct depth information only from (left) by decoding uniqueness codewords.

The experiment is conducted with a plastic doll which contains plentiful colors. Fig. 4 shows the test target, and its image with pattern illumination. From the image we can see that, the green color of the surface greatly polluted the project pattern elements, especially to the green color pattern elements. Learning curve of the model without pairwise features $f_{i,j}^{P}(C_i,C_j)$ and scene color is displayed by Fig. 5. It shows a softmax function as a multi-classifier with respect to single pattern element without considering the dependences and causal effect. With the training model fed with enough label data, the training error and test error become similarly. That shows that pure multi-classifier model is underfitting.

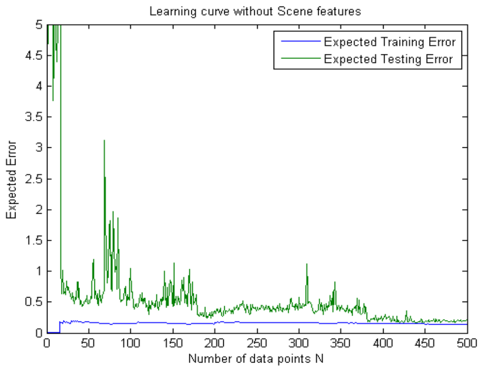

Fig. 5, Learning curve of the model without consideration of the effect from surface background colors.

Fig. 6 shows the learning curve by the proposed unified graphical model based method. It is enriched with factors like local correlations and causal effect due to surface color. It converges quickly and achieves high performance in both training and testing dataset. And almost all the color pattern element within colourful background regions can be correctly identified. To show the final result visually, the decoded features are triangulated and the rendered 3D reconstruction results are displayed in Fig. 7.

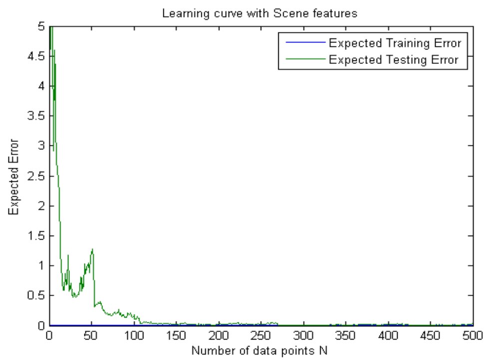

Fig. 6. Learning curve of proposed graphical model with the considerations of pairwise factors and causal factor.

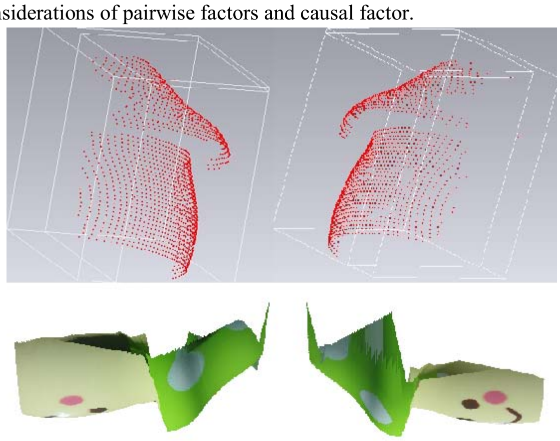

Fig. 7. 3D reconstruction of the target based upon the decoding results. The point cloud model (up), the rendered 3D models (bottom).

IV. Conclusion and Future Work

This paper presents a novel and practical method for the decoding of a color pseudorandom structured light pattern. In the label procedure, the color pattern is decomposed into different binary pattern images, and thus makes the label procedure more easily and efficiently. A unified probabilistic graphical model is constructed to encode uniqueness in local pattern. The CRF is utilized to capture dependences between pattern elements in one codeword. And the BN method is used to capture the causal effect from object color. By combining the CRF and BN approaches, we can get a local probabilistic model of unique pattern, which is supervised, locally and robust to complex color objects. A colorful surface is experimented to demonstrate its feasibility. And the 3D reconstruction results are also provided to show its robustness.

In the future, the model can be further improved in cooperate with a latent variable of depth or a smoothness variable to well treat with the surface discontinuities or occlusions.

References

[1] J. Salvi, J. Pages, and J. Batlle, “Pattern codification strategies in structured light systems,” *Pattern Recognition*, vol. 37, no. 4, pp. 827-849, 2004.

[2] C. Albitar, P. Graebling, and C. Doignon, “Robust structured light coding for 3D reconstruction computer vision,” in *Proc. IEEE Int’l Conf. Computer Vision*, pp. 1-6, 2007.

[3] J. Salvi, J. Batlle, and E. Mouaddib, “A robust-coded pattern projection for dynamic 3D scene measurement,” *Pattern Recognition Letters*, vol. 19, pp. 1055-1065, 1998.

[4] Z. Song and R. Chung, “Determine both surface position and orientation in structured-light-based sensing,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 32, no.10, pp. 1770-1780, Oct. 2010.

[5] C. Schmalz, and E. Angelopoulou, “A graph-based approach for robust single-shot structured light.” In *Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 80-87, 2010.

[6] L. Zhang, B. Curless, and S. Seitz, “Rapid shape acquisition using color structured light and multi-pass dynamic programming,” in *Proc. First Int’l Symp. 3D Data Processing, Visualization, and Transmission*, pp. 22-36, 2002.

[7] A. O. Ulusoy, F. Calakli, and G.Taubin, “Robust one-shot 3D scanning using loopy belief propagation.” in *Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 15-22, 2010.

[8] X. Zhang, Y. Li, and L. Zhu, “Color code identification in coded structured light.” *Applied Optics* vol. 51, no. 22, pp. 5340-5356, 2012

[9] P. Fechteler, and P. Eisert, “Adaptive color classification for structured light systems.” *Computer Vision, IET* 3.2, pp. 49-59, 2009.

[10] L. Zhang and Q. Ji, “Image segmentation with a unified graphical model,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 32, no. 8, pp. 1406-1425, Aug. 2010.

[11] J. Shi and J. Malik, “Normalized cuts and image segmentation,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888-905, Aug. 2000.

[12] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques*. The MIT Press, 2009.

[13] Z. Zhang, “A flexible new technique for camera calibration,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, Nov. 2000.