Submitted to *Sensors*. Pages 1 - 17.

OPEN ACCESS

Article

Adaptive Color Calibration Based One-shot Structured Light System

Yu Zhou *, Dongwei Zhao, Yao Yu, Jie Yuan and Sidan Du*

School of Electronic Science and Engineering, Nanjing University, Nanjing 210093, Jiangsu, China; E-Mail: nackzhou@nju.edu.cn

* Author to whom correspondence should be addressed; E-Mails: nackzhou@nju.edu.cn (Y.Z.); coff128@nju.edu.cn (D.S.D.); Tel.: +86-25-8359-7453; Fax: +86-25-8359-7453.

Version July 21, 2012 submitted to Sensors. *Typeset by LaTeX using class file mdpi.cls*

Abstract: In one-shot color structured light system, the color of stripe patterns is usually distorted with respect to color crosstalk, ambient light and albedo of the scanned objects, leading to mismatch in the correspondence of color stripes between the projected and captured images. In this paper, an adaptive color calibration as well as a Discrete Trend Transform algorithm is presented to achieve high-resolution 3D reconstructions. The adaptive color calibration, according to the relative albedo in RGB channels, can improve the accuracy of labeling stripe by alleviating the effect of albedo and ambient light while decoding the color. Furthermore, the Discrete Trend Transform in M channel makes it effective to detect weak stripes due to the uneven surface and reflectance characteristics of the scanned objects. With this approach, the presented system is suitable for scanning the moving objects and generating high-resolution 3D reconstructions without the need of dark laboratory environments.

Keywords: three-dimensional image acquisition; adaptive color calibration; discrete trend transform; structured light.

1. Introduction

3D shape acquisition has received considerable research interest in the last decade. There are some groups of techniques for 3D shape acquisition, such as stereo vision, Time Of Flight(TOF) and structured light[1]. Compared with TOF and stereo vision techniques, the main advantages of structured

light techniques include the easy image processing involved and the high accuracy achieved in the 3D reconstruction[2]. Structured light techniques are based on triangle principle. In the first place, some designed patterns are projected onto the scanned objects, and a camera captures the scene. Secondly, the correspondence between projected patterns and captured patterns is established. Finally, the 3D coordinates of the objects are derived from correspondence between projected pattern and detected pattern combining geometric calibration parameter of the projector and the camera.

In the development of structured light techniques, the time-multiplexing techniques, also called multi-shots techniques, is the first to appear. These techniques can achieve high accuracy and resolution through encoded points by projecting a sequence of patterns. However, there is a trade-off between the scanning speed and the resolution. In the recent years, fast time-multiplexing techniques have been developed and make it possible to acquire the shape of moving object, but such techniques need a specialized hardware. To achieve high resolution and speed at the same time, authors in[3], [4] propose a one-shot structured light system for rapid range acquisition. For one-shot structured light systems, the final reconstruction result is significantly affected by three vital processes: the designation of projected patterns, detection of coded cell in captured image and establishment of correspondence between projected and captured coded cells. In recent years, with more research and efforts have been made, this technique is generating higher quality 3D reconstruction result, making it quite useful and widespread in industry and cultural heritage.

The projected pattern has a strong influence on 3D shape acquisition. Obviously the pattern itself should be easy to detect and possible to establish correspondence between projected and captured coded cells. Different choices of patterns can be found in [5],[6]. Examples are the use of color stripe[7], a unique shape[8], modulated patterns[9] and color coded grids[10]. In [11], a 2D M-array pattern with a Hamming distance is suggested for robustness consideration. The matrix is comprised of many symbols and this technique enhances the tolerance of failure to recover symbols. In this pattern each symbol contains several pixels and thus the reconstruction resolution is limited. In [12], authors propose a self-adaptive system for real-time range acquisition. Pattern color, geometry, tracking and graph cut are used to solve the corresponding problem. By using coded color, the system can obtain denser result than that using shape or modulated patterns. However, some factors, such as reflectance characteristics of scanned object, ambiguity due to uneven surface and ambient light, have great effect upon color classification and detection of coded cells. To avoid the influence of these factors, in the work by Zhang et al.[13], a method using multi-pass dynamic programming and edge-based reconstruction is presented. This method alleviates albedo influe, but the bad side is that sub-pixel-accuracy position of the edge can not be located by edge-based reconstruction method. Pages et al.[2] propose a peak-based coding strategy which can improve the resolution without loss of accuracy, however the albedo of illuminated object is modeled by a static matrix, which limits the usage of application to static scene. For the purpose of robustness under ambient light, Benveniste et al.[14] develop a structured light system which decodes the color using a color invariant and optimize the projected patterns by flexibly changing stripe color for different colored objects.

Addressing the impact of these factors, in this paper, we focus on one-shot color structured light system and propose an adaptive color calibration as well as a Discrete Trend Transform (DTT) algorithm to obtain high-resolution 3D point cloud without the need of dark laboratory environments. Firstly,

a relative albedo function between two color channels is proposed to calibrate color adaptively. With this method, the distorted color caused by different albedo in RGB channels is calibrated adaptively. Meanwhile, white ambient light in the scene can be canceled during color classification. Secondly, considering the different intensity caused by uneven surface or viewpoint, a novel M channel, DTT and sub-pixel peak localization algorithm are proposed to segment and detect the stripes. These techniques make significant improvement in accuracy of locating and labeling the stripes. In the literature, the goal of Fechteler's work is most similar to our work, where the KMeans and clustering are utilized to classify the color adaptively[16].

2. System Setup and Framework

To realize high-resolution 3D reconstruction, a DLP projector with a WUXGA resolution (1920x1080) is used to project structured color light, and an camera Nikon D200 with a resolution of 5.2 mega pixels (2560x2048) is used to grab the scene. Higher resolution of camera and projector means denser and more accurate point cloud can be achieved. The relative direction angle between the projector and camera are $17°$.

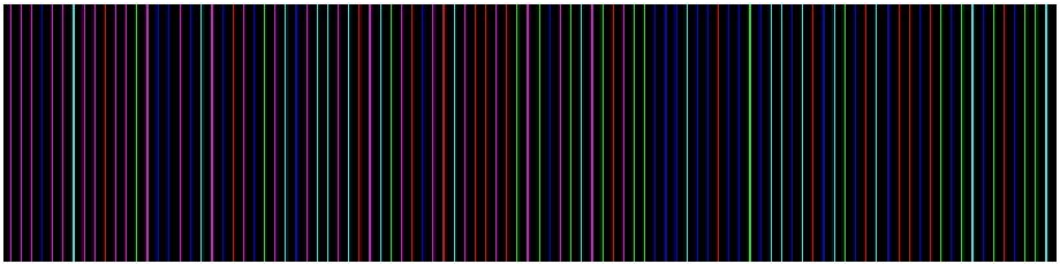
In order to acquire 3D shape in a natural environment instead of dark environment, the following steps are performed:

1. Project stripe patterns onto the surface of target objects, and capture an image under ambient light.
2. Calibrate the geometric distortion of camera and projector. We refer to the method proposed by Pages[2].
3. Adjust color by applying static and adaptive color calibration. The color crosstalk is approximated in static color calibration, and then the adaptive color calibration is adopted to adjust color with respect to object's color and ambient light.
4. Segment the stripes and locate the peak of stripes. The proposed DTT algorithm is applied to stripe segmentation. Furthermore, the sub-pixel accurate localizations of peaks are derived.
5. Classify the color of stripes and figure out the correspondence between projected patterns and captured patterns by dynamic programming algorithm[17] with a cost function of hue value.
6. Generate the 3D point clouds and reconstruct the surfaces of objects through 3D delaunay triangulation.

2.1. Projected Coding Patterns

We use a modified De Bruijn sequence as the arrangement of color stripes to design projected patterns. Supposing $D(k, n)$ denotes a k-array, n-order De Bruijn sequence, the element would be $E = \{0, 1, 2, 3, 4, 5\}$. Each element is assigned a color with a distance of $60°$ in hue value, as listed in Table 1. We consider that neighbor stripe should differ at least two channels in (R, G, B). With this constraint, only three nodes are permitted to follow one certain node. Take element 0 for example, the

Figure 1. The cut-out of generated stripe patterns.



possible next following node should only be $1, 2$ or 3. By this way, a De Bruijn sequence with a length of 162 is generated. Figure 1 is the cut-out of generated stripe patterns.

Table 1. Color assignment for each element.

element	**color**(R, G, B)	**hue value**
0	$red(255, 0, 0)$	$0°$
1	$green(0, 255, 0)$	$60°$
2	$blue(0, 0, 255)$	$120°$
3	$cyan(0, 255, 255)$	$180°$
4	$magenta(255, 0, 255)$	$240°$
5	$yellow(255, 255, 0)$	$300°$

3. Static and Adaptive Color Calibration

The color structured light system in a robust manner consists of two key processes: the recovery of color in grabbed image in contrast with the projected color on the object's surface, and the establishment of correspondence between the projected patterns and captured patterns correctly. The distortion of color is due to the following reasons:

- The color crosstalk between the projector and camera.
- The ambient light in the environment.
- The different colored object results in different albedo of RGB channels.
- The stripes in captured image usually vary in both amplitude and width, caused by the uneven surface or reflectance characteristics of the scanned objects.

3.1. Static Color Calibration

The color distortion caused by projector-camera color crosstalk is adjusted by a static color calibration, in which the parameters are stable for a pairwise projector-camera and measured in advance.

The camera captures the reflected light through RGB channel sensors as an image. The model of this process is formulated as Equation (1) by Caspi et al.[18].

$$\underbrace{\begin{bmatrix} c^r \\ c^g \\ c^b \end{bmatrix}}_{I_c} = \underbrace{\begin{bmatrix} x_{rr} & x_{rg} & x_{rb} \\ x_{gr} & x_{gg} & x_{gb} \\ x_{br} & x_{bg} & x_{bb} \end{bmatrix}}_{X} \underbrace{\begin{bmatrix} a^r & 0 & 0 \\ 0 & a^g & 0 \\ 0 & 0 & a^b \end{bmatrix}}_{A} \Big(\underbrace{\begin{bmatrix} p^r \\ p^g \\ p^b \end{bmatrix}}_{I_p} + \underbrace{\begin{bmatrix} o^r \\ o^g \\ o^b \end{bmatrix}}_{O} \Big) \tag{1}$$

where I_c is the observed color through camera and I_p denotes the corresponding projected color projected by a projector. O addresses the ambient light illumination. X indicates the projector-camera color crosstalk matrix and A represents albedo matrix of the object surface. During the process of static calibration, the color crosstalk matrix X in Equation (1) is approximated by projecting solid color stripe patterns to a white planar board.

3.2. *Adaptive Color Calibration*

In order to recover the stripe's real color and cancel the influence of different albedo in RGB channels and ambient light in the scene, the relative albedo is then defined to calibrate stripe color adaptively, as shown in Equation (2).

$$A = \begin{bmatrix} a^r & 0 & 0 \\ 0 & \alpha a^r & 0 \\ 0 & 0 & \beta a^r \end{bmatrix} = a^r \begin{bmatrix} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \beta \end{bmatrix} \tag{2}$$

where a^r is the red channel reflectance ratio, α indicates the green channel relative reflectance ratio compared to red channel, similarly, β denotes the blue channel relative reflectance ratio.

In processes of stripe segmentation and color classification, it is best to use I_p as an input. Unfortunately, I_p is distorted by crosstalk effect and albedo matrix. Given camera output I_c, we compute calibrated color through color calibration processes, which can be expressed as Equation (3).

$$\begin{aligned} I_p + O &= \tilde{A}^{-1} X^{-1} I_c \\ &= a^{-r} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \tilde{\alpha} & 0 \\ 0 & 0 & \tilde{\beta} \end{bmatrix}^{-1} X^{-1} I_c \end{aligned} \tag{3}$$

in which $\tilde{A}$, compromised of $\tilde{\alpha}$ and $\tilde{\beta}$, is the estimation of A. Then the calibrated color I_a, defined as Equation (4), is the input for stripe segmentation and color classification.

$$\begin{aligned} I_a &= a^r (I_p + O) \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \tilde{\alpha} & 0 \\ 0 & 0 & \tilde{\beta} \end{bmatrix}^{-1} X^{-1} I_c \end{aligned} \tag{4}$$

The color crosstalk matrix X can be obtained from static color calibration. In the following section we conclude that $\tilde{\alpha}$ and $\tilde{\beta}$ can be estimated from relative albedo estimation. On the second hand, influence of ambient light O and a^r can both be canceled in color classification and stripe segmentation.

3.3. Relative Albedo Estimation

We estimate $\tilde{\alpha}$ by define a relative albedo function between red and green channel, similarly, $\tilde{\beta}$ is estimated by comparing red and blue channel. A histogram for each channel is produced. For example, the histogram of red channel is expressed as Equation (5):

$$H^r = \{(H_1^r, W_1^r), (H_2^r, W_2^r), \cdots, (H_n^r, W_n^r)\} \tag{5}$$

where $H_i^r(1 \leq i \leq n)$ is the histogram bin and $W_i^r(1 \leq i \leq n)$ is the bin value, the superscript r, g and b denote the red, green and blue channel respectively. To match H^r, H^g is transformed into $H^{g\prime}$, and a flow matrix $\mathbf{f} = \{f_{ij}\}(1 \leq i \leq n, 1 \leq j \leq n)$ is defined to represent the transition process. An specified $\mathbf{f}$ minimizes overall cost function and the cost function is indicated as Equation (6).

$$cost(H^g, H^r) = \sum_{i=1}^{n}\sum_{j=1}^{n} |j-i| f_{ij} \tag{6}$$

where f_{ij} denotes the flow from H_i^g to H_j^r, and $|j-i|$ denotes the flow distance, which implies that we encourage a flow from one bin in the histogram to another bin with shorter distance. Equations from Equation (7) to Equation (10) are the constraints of the process.

$$f_{ij} \geq 0 \text{ for } 1 \leq i \leq n, 1 \leq j \leq n \tag{7}$$

$$\sum_{i=1}^{n} f_{ij} \leq W_j^r, 1 \leq j \leq n \tag{8}$$

$$\sum_{j=1}^{n} f_{ij} \leq W_i^g, 1 \leq j \leq n \tag{9}$$

$$\sum_{i=1}^{n}\sum_{j=1}^{n} f_{ij} = \min(\sum_{i=1}^{n} W_i^g, \sum_{j=1}^{n} W_j^r) \tag{10}$$

By adopting the Earth Mover's Distance (EMD) algorithm[15], which measures the least amount work needed to match between two histograms through liner programming, a flow matrix $\mathbf{f} = \{f_{ij}\}$ from one histogram to another can be obtained. The f_{ij} is pixel number transited from bin i in one histogram to bin j in another. Then we propose a relative albedo function between green and red channel as Equation (11) and obtain the estimated relative reflectance ratio $\tilde{\alpha}$(i) for each grey level.

$$\tilde{\alpha}(i) = \frac{(\sum_{j=1}^{n} f_{ij})i}{\sum_{i=1}^{n}\sum_{j=1}^{n}(f_{ij}j)}, 1 \leq i \leq n, 1 \leq j \leq n \tag{11}$$

Relative reflectance ratio $\tilde{\beta}$(i) can be estimated similarly. After calibration, I_a serves as the input of stripe segmentation and color classification.

4. Stripe Segmentation and Peak Localization

After static and adaptive color calibration, the color of stripe can be classified correctly in the following Section 5. But the pixel-wise O and red channel reflectance ratio a^r still remain in I_a, which make effective stripe segmentation and accurate peak localization challenging. So we propose a novel M channel and DTT algorithm to derive robust stripe segmentation result.

4.1. M Channel Definition

An M channel, which is a function of RGB channels in I_a, is proposed as Equation (12) to suppress ambient light O .

$$M_{ij} = \max\left(C^r_{ij}, C^g_{ij}, C^b_{ij}\right) - \min\left(C^r_{ij}, C^g_{ij}, C^b_{ij}\right) \tag{12}$$

where C^r_{ij} in Equation (13) is the red channel value of pixel (i, j) in I_a, and the superscript r, g and b are the red, green and blue channel respectively.

$$\begin{pmatrix} C^r_{ij} \\ C^g_{ij} \\ C^b_{ij} \end{pmatrix} = a^r \begin{pmatrix} p^r_{ij} + o^r_{ij} \\ p^g_{ij} + o^g_{ij} \\ p^b_{ij} + o^b_{ij} \end{pmatrix} \tag{13}$$

Given the assumption that ambient light is mostly white light, i.e., $o^r \approx o^g \approx o^b$, the M channel can be simplified as Equation (14).

$$\begin{aligned} M_{ij} &= a^r \max \begin{pmatrix} p^r_{ij} + o^r_{ij} \\ p^g_{ij} + o^g_{ij} \\ p^b_{ij} + o^b_{ij} \end{pmatrix} - a^r \min \begin{pmatrix} p^r_{ij} + o^r_{ij} \\ p^g_{ij} + o^g_{ij} \\ p^b_{ij} + o^b_{ij} \end{pmatrix} \\ &= a^r \max \begin{pmatrix} p^r_{ij} \\ p^g_{ij} \\ p^b_{ij} \end{pmatrix} - a^r \min \begin{pmatrix} p^r_{ij} \\ p^g_{ij} \\ p^b_{ij} \end{pmatrix} \\ &= a^r \max \begin{pmatrix} p^r_{ij} \\ p^g_{ij} \\ p^b_{ij} \end{pmatrix} \end{aligned} \tag{14}$$

Note that there is at least one invalid color channel in designed stripe patterns, therefore the $\min(p^r_{ij}, p^g_{ij}, p^b_{ij})$ is always zero. As a result, the M channel suppresses white ambient light. But the red channel reflectance ratio a^r is still an interference.

The reason for introducing M channel is for stripe segmentation. To achieve accurate stripe segmentation result, we just need to focus on the improved amount of intensity instead of the intensity itself. M channel describes the amount of improved illumination intensity perfectly. Another advantage of M channel is that while keeping original characters it integrates RGB color channels into one channel, which makes it convenient for further process. Figure 2 is the M channel of Figure 8(e).

Figure 2. An example of M channel

4.2. Stripe Segmentation

The stripe pattern is vertical, we process M channel horizontally. A stripe in captured image means a peak in M channel. Instead of segmenting stripes by illuminance value, we detect stripes by locating intensity rising edge and falling edge in M channel. The rising edge and falling edge information is derived by applying DTT in Equation (15) and Equation (16).

$$T_{ij} = \sum_{k=h+1}^{j+N} \sum_{h=j}^{j+N-1} sign(M_{ik} - M_{ih}) \tag{15}$$

and

$$sign(a) = \begin{cases} 1 & a > 0 \\ 0 & a = 0 \\ -1 & a < 0 \end{cases} \tag{16}$$

where T_{ij} indicates the local trend of pixel (i, j) in M channel. The local trend measures the probability of increasing or decreasing trend in a given window. The size of window is controlled by N, and N is usually assigned a value less than half of the stripe width in captured image.

In Equation (15), obviously $|k - h| \leq N$. Since N is a small number, assume that reflectance ratio is locally continuous, i.e., $a^r_{ik} \approx a^r_{ih}$, then we get Equation (17).

$$\begin{aligned} & sign(M_{ik} - M_{ih}) \\ = \; & sign(a^r_{ik} \max \begin{pmatrix} p^r_{ik} \\ p^g_{ik} \\ p^b_{ik} \end{pmatrix} - a^r_{ih} \max \begin{pmatrix} p^r_{ih} \\ p^g_{ih} \\ p^b_{ih} \end{pmatrix}) \\ = \; & sign(\max(\mathbf{p}_{ik}) - \max(\mathbf{p}_{ih})) \end{aligned} \tag{17}$$

where $\mathbf{p}_{ik} = (p^r_{ik}, p^g_{ik}, p^b_{ik})^T$ and $\mathbf{p}_{ih} = (p^r_{ih}, p^g_{ih}, p^b_{ih})^T$. Thus Equation (15) can be rewritten to Equation (18), and therefore the T_{ij} is only related to projected pattern and is independent of reflectance ratio and ambient light.

$$T_{ij} = \sum_{k=h+1}^{j+N} \sum_{h=j}^{j+N-1} sign(\max(\mathbf{p}_{ik}) - \max(\mathbf{p}_{ih})) \tag{18}$$

The DTT in Equation (18) works in the following way:

- at falling edge in M channel, T arrives its maximum $\frac{N(N+1)}{2}$.
- at rising edge in M channel, T arrives its minimum $-\frac{N(N+1)}{2}$.
- otherwise value of T is between maximum and minimum.

Figure 3. Robustness of local trend. Four peaks in $f(n)$ conform to Gaussian distribution and have different arguments(stripe width and amplitude), but local trend T has the same Max-to-Min transition pattern at each peak position.

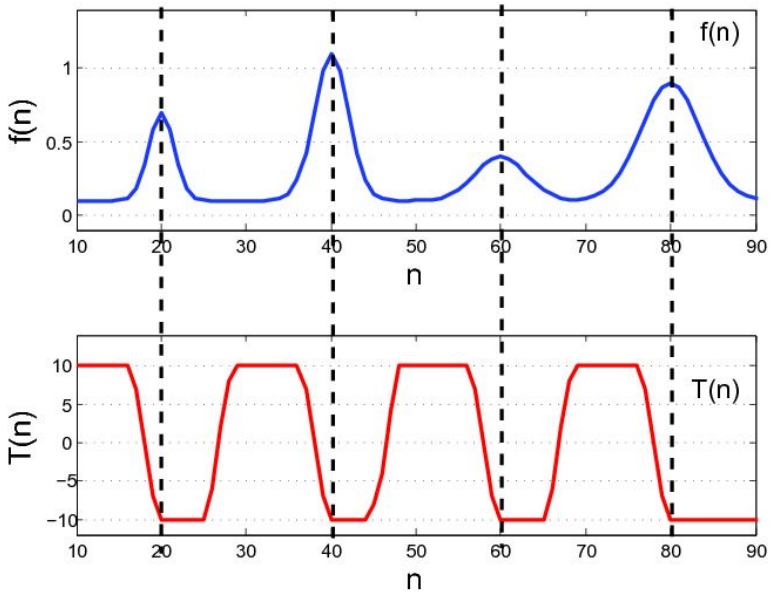
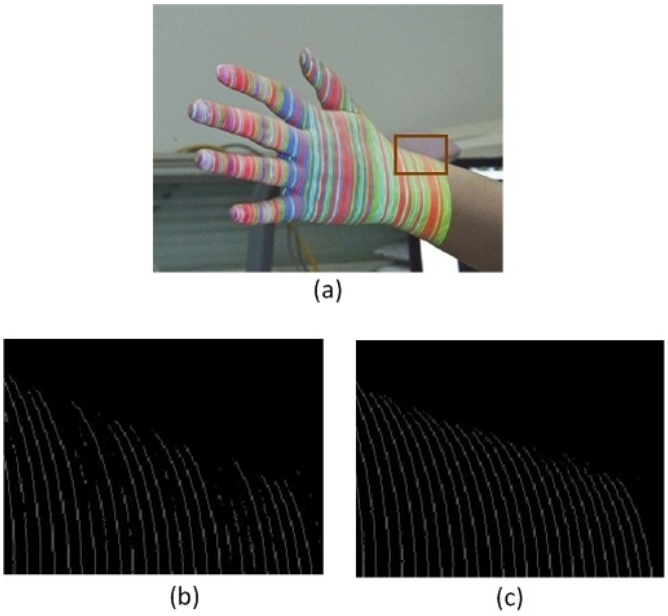


Figure 3 illuminates how DTT works. For example, $f(n)$ is a scan-line in M channel. If $f(n)$ is locally monotonically decreasing in the window, then T arrives at maximum value $\frac{N(N+1)}{2}$. On the contrary, T arrives at its minimum value $-\frac{N(N+1)}{2}$ at $f(n)$'s rising edge. Values between maximum and minimum denote that $f(n)$ is neither monotonically increasing nor decreasing. The transition from maximum value to minimum value in T indicates a peak in M channel. So Max-to-Min transition in T is searched to locate stripes. The rising edge in M channel is labeled as the start of a stripe area, and the falling edge in M channel is labeled as the end of a stripe area. The result of peak detection using DTT illustrated in Figure 4. We can figure out that DTT is more robust than local adaptive threshold method[19] in term of weak stripes.

Actually, with the help of M channel and DTT, there is no need to cut the foreground from background. The standard of DTT detecting color stripes is a pair of strict rising edge and strict falling edge in M channel. The standard is so strict that very few points in background is taken as candidate stripes after DTT. Robustness is enhanced by excluding the influence of backgrounds.

Figure 4. A comparison of stripe segmentation result. (a) captured source image, (b) stripe segmentation result using local adaptive thresholding method, (c)stripe segmentation result using DTT.



4.3. Peak Localization

Accurate sub-pixel peak localization needs to be estimated for the derivation of 3D point cloud. Some existing sub-pixel peak localization algorithms are detailed and compared in [20]. In order to estimate accurate peak position, the maximum value M_{max} of M channel is searched in each stripe area and its position is labeled as I_{max}. The estimated sub-pixel peak position $I_{estimated}$ is proposed as Equation (19):

$$I_{estimated} = \frac{\sum I_i \cdot M_i}{\sum I_i}, \text{ for } M_i \geq \alpha M_{max} \tag{19}$$

where I_i is the horizonal offset, M_i is the M channel intensity value, and α is the related ratio, which defines the pixels related to M_{max} around I_{max}. In the following experiments α is set to 0.8.

5. Color Classification

The color is classified in hue space. Assuming ambient light is mostly white light and taking the color channel is $C^r > C^g > C^b$ as an example, we calculate hue value from Equation (20).

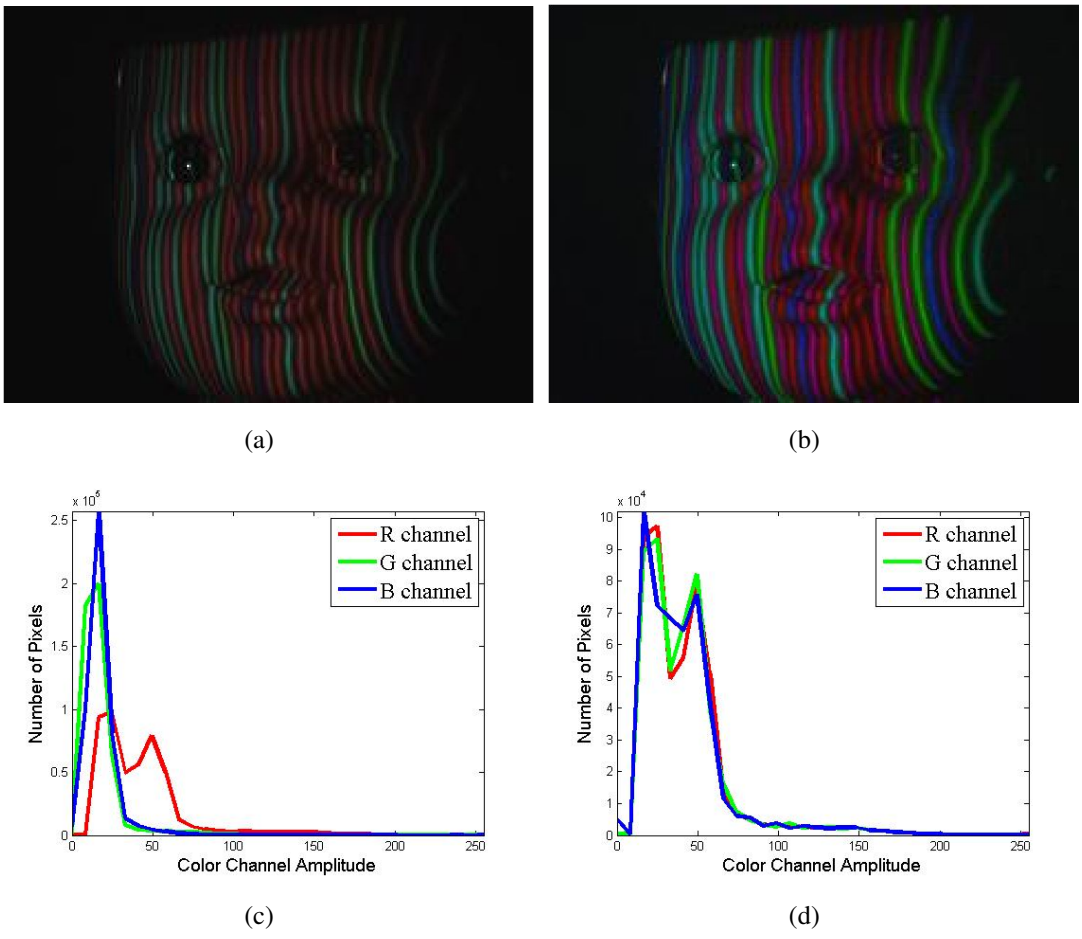
$$\begin{aligned} h &= \frac{C^g - C^b}{C^r - C^b} \cdot \frac{\pi}{3} \\ &= \frac{a^r(p^g + o^g) - a^r(p^b + o^b)}{a^r(p^r + o^r) - a^r(p^b + o^b)} \cdot \frac{\pi}{3} \\ &= \frac{p^g - p^b}{p^r - p^b} \cdot \frac{\pi}{3} \end{aligned} \tag{20}$$

Where (C^r, C^g, C^b) is the color intensity in I_a. h in other cases can be calculated by equations similar to Equation (20). Equation (20) demonstrates that the O and a^r are canceled, and the h of colors

in I_a is only related to projected patterns, thus the classification process is independent of reflectance characteristics of the scanned objects and white ambient light.

The correspondence between detected pattern and projected pattern is founded by applying dynamic programming[17]. We compare three detected neighboring stripes each with three projected neighboring stripes. The sum of hue value difference is defined as the score function. By considering neighboring conditions we correctly handle the edge conditions in cases where occlusion occurs.

Figure 5. An example of adaptive albedo calibration. (a) face model under illumination patterns, (b) calibrated image of face model, (c), (d)histograms of(a), (b) respectively.

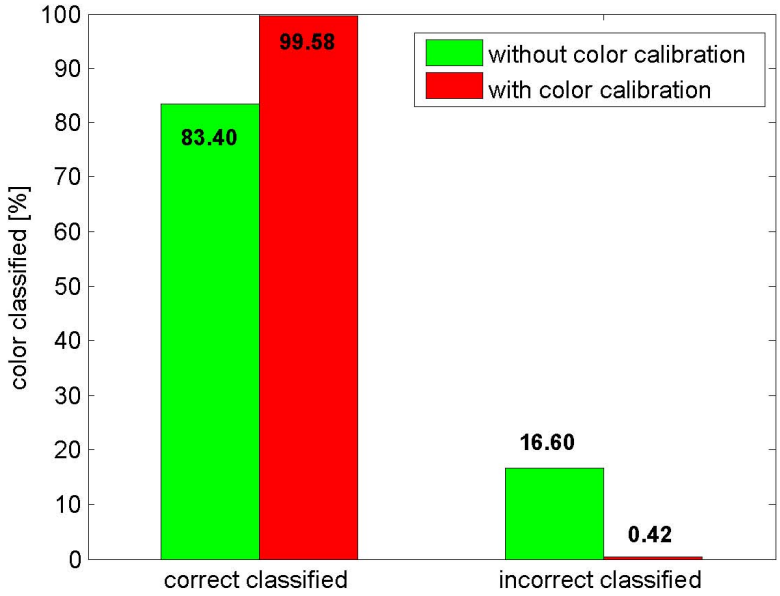


6. Experiments and Discussions

6.1. Adaptive Color Calibration Performance

Adaptive color calibration is a crucial method to assure the accuracy rate of color classification. This method can adjust stripe's color adaptively with respect to the different albedo in RGB channels. Figures 5(a) and 5(c) illustrate that the gain of R channel is stronger than the G and B channel. After adaptive color calibration, the gains of three channels are almost at same level in Figures 5(b) and 5(d). Some quantitative comparison between the adaptive albedo calibration and non-adaptive calibration have

Figure 6. A comparison of color classification results.



been done. The results in Figure 6 is clearly visible that adaptive albedo calibration increases the number of correctly labeled points (more than 15%).

6.2. *Peak Localization Performance*

In this subsection, we mainly focus on analyzing the error of estimated sub-pixel peak position. The performance of the traditional methods are compared with our method referred as Max-Min Weighted Average Method (**MMWA**). Some traditional methods are listed as follows:

- *Max Method*(**MAX**)[2]. Pages et al. use maximum intensity value to define M channel and choose localizations where M channel reaches maximum value as estimated peak position.

$$I_i = max(R_i, G_i, B_i) \tag{21}$$

$$center = x \text{ where } I_x = max(I_i)$$

- *Weighted Average Method*(**WA**)[21]. It calculates the average intensity value of RGB channels and then uses a weighted algorithm in whole area of stripe to derive the peak position.

- *Midpoint Method*(**MID**)[12]. It simply uses the midpoint of a stripe as feature point.

- *Probability Method*(**PM**)[16]. Fechteler et al. use probability method to estimate peak localization. They detect maximum in each RGB channel and assign each color peak a probability of being a valid stripe. Peak localization is estimated by these probabilities.

$$center = \frac{C_1 * P_1 + C_2 * P_2}{P_1 + P_2} \tag{22}$$

where P_i is the probability of C_i being a valid stripe

To simulate captured stripes, we generate a designed color stripe patterns. As shown in Equation(23), intensity of a valid color channel is consistent with Gaussian Distribution and corrupted with noise. Each RGB channel is additionally polluted by some offset o simulating ambient light. For example, a red stripe compromises a valid red channel generated by Equation(23) while invalid Green and Blue channels contain offset o only.

$$I(n, c, \sigma, \beta, A, o) = Ae^{-\frac{(n-c)^2}{2\sigma^2}} + \beta A\epsilon + o \tag{23}$$

where A denotes the amplitude of color intensity. n is the measured pixel. βA is the noise amplitude and $\epsilon \in (0, 1)$. We consider different noise levels: SNR = 25dB, 20dB, 18dB. c is the stripe peak position. σ controls the stripe width. o is the intensity offset.

RMS error of peak localization c, defined as Equation (24) , is measured in each method by analyzing 10000 samples. The average RMS Error at different noise levels are listed in Table 2, as σ changes from 0.3 to 0.6.

$$RMS\_Error = \sqrt{\frac{\sum (c_i - c_i')^2}{N}} \tag{24}$$

In all methods, RMS error increases as noise increases from $25dB$ to $18dB$, while MMWA derives least RMS error at the same noise conditions. Figure 7 depicts that WA and MID are sensitive to light offset, nevertheless, our method performs high accuracy even in a strong ambient light environment, this is because the offset is suppressed when deriving M channel. Form Table 2 and Figure 7, we can conclude that our method outperforms other methods in terms of RMS Error with respect to noise, stripe width and ambient light.

Figure 7. RMS Error versus o/A (SNR=18dB,$\sigma = 0.3$).

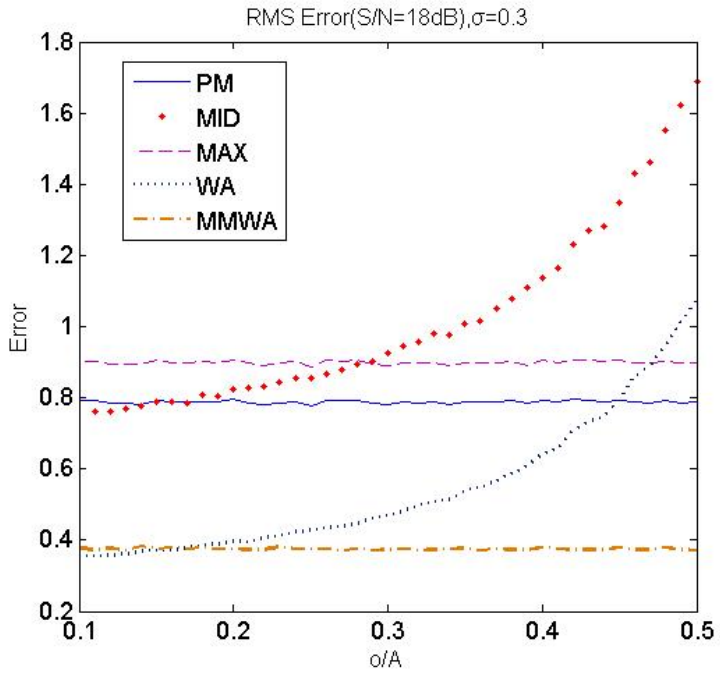
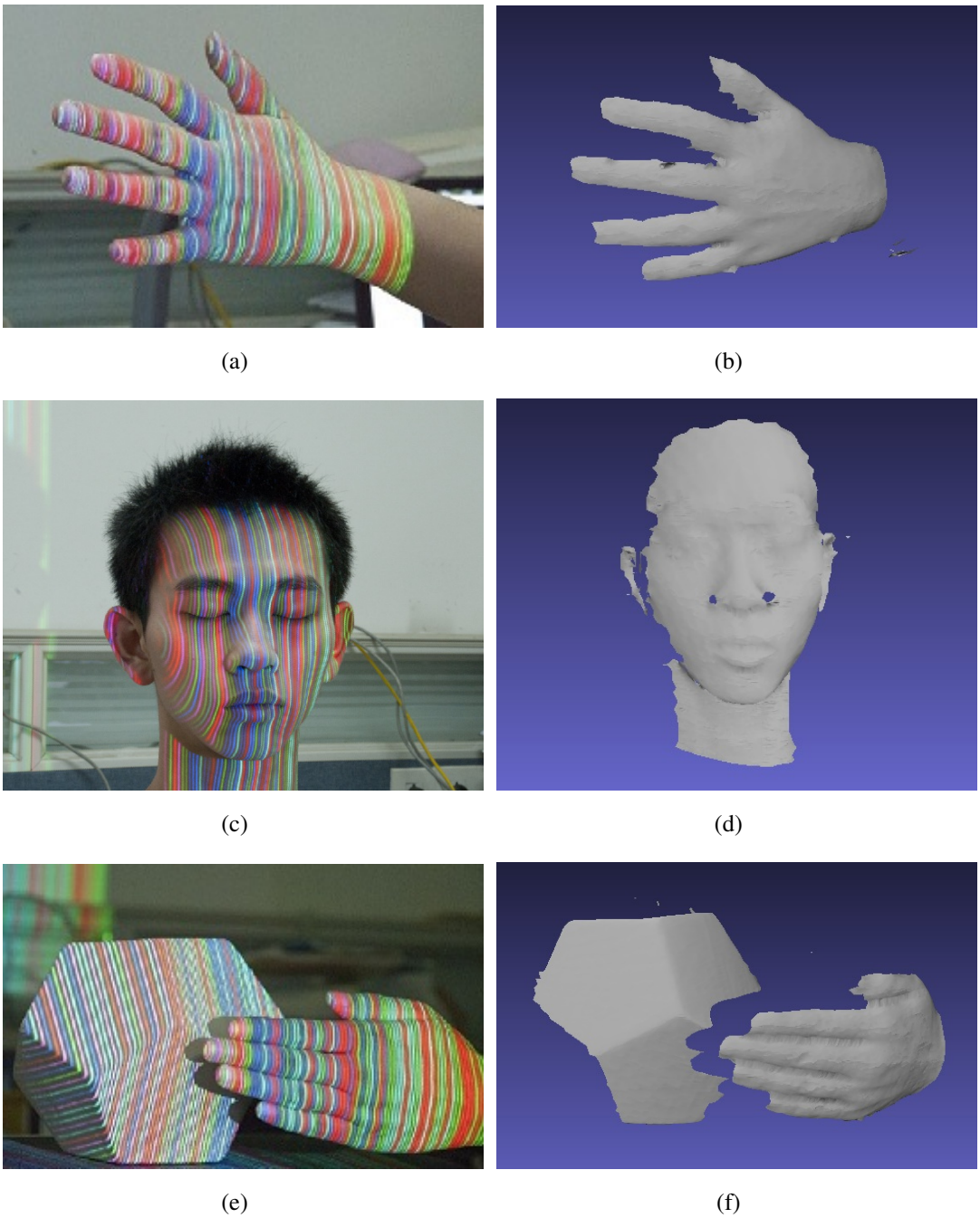


Figure 8. Reconstruction examples: (a) hand, (c) face and (e) hand with background of hexahedron under the pattern illumination. Reconstructed surface of the (b) hand, (d) face and (f) hand with background of hexahedron.



(a) (b)

(c) (d)

(e) (f)

Table 2. Average RMS Error at different noise.

SNR	MMWA	PM	MID	MAX	WA
25dB	0.27	0.75	0.89	0.87	0.38
20dB	0.46	1.05	1.04	1.14	0.54
18dB	0.62	1.21	1.12	1.28	0.61

6.3. Reconstruction Results

Figures 8(a) and 8(c) demonstrate the hand and face are projected with stripe patterns in a natural environment. With the presented methods, the color of stripes could be classified correctly, the weak stripes could be detected effectively and the surfaces of object could be reconstructed robustly even being illuminated by such a strong ambient light, depicted in Figure 8(b) and 8(d). Meanwhile, the reconstructed surfaces of hand and face illustrate the richness in detail. Figures 8(b) and 8(d) contain 85687 and 75371 vertices respectively. Note that there is a hole nearby the nose in Figure 8(d), due to the occlusion resulted in the captured image. Furthermore, Figures 8(e) and 8(f) demonstrate that the 3D shapes of multiple objects with different albedo in a same scene could be reconstructed simultaneously, because color was calibrated adaptively according to the gray level and the stripe are segmented by DTT.

7. Conclusion

We have presented a color structured light system for robust 3D shape acquisition with regard to reflectance characteristics of scanned object, ambiguity due to uneven surface as well as white ambient light. Our contributions lie in two aspects of the proposed approach. The first is a novel method for calibrating color adaptively according to the colored objects and white ambient light in the scene, as Section 3.2 and 3.3 discussed, thus enhancing the robustness of the system and widening the range of potential applications. The second contribution lies in the effectiveness of finding weak stripe caused by uneven surface of scanned object, by using using M channel and DTT algorithm. Furthermore, we propose an algorithm to locate the sub-pixel peak of stripes. Through some experimental evaluations, we demonstrate that this structured light system employing the proposed techniques can get high-resolution 3D reconstructions without the need of dark laboratory environments.

Acknowledgements

This work was partially supported by Grant No. BK2010391, BK2011563 from Natural Science Foundation of Jiangsu Province and Grant No. 61100111 from Natural Science Foundation of China.

References

1. Sansoni, G.; Trebeschi, M.; Docchio, F. State-of-The-Art and Applications of 3D Imaging Sensors in Industry, Cultural Heritage, Medicine, and Criminal Investigation. *Sensors* **2009**, *9*, 568–601.

2. Pages, J.; Salvi, J.; Collewet, C.; Forest, J. Optimised De Bruijn patterns for one-shot shape acquisition. *Image Vision Comput.* **2005**, *23*, 707–720.
3. Boyer, K.L.; Kak, A.C. Color-encoded structured light for rapid active ranging. *IEEE Trans. Pattern Anal. Mach. Intell.* **1987**, *1*, 14–28.
4. Proesmans, M.; Van Gool,L.J. One-shot active 3D image capture. *Proc. of SPIE* **1997**, *3023*, 50.
5. Salvi, J.; Pages, J.; Batlle, J. Pattern codification strategies in structured light systems. *Pattern Recogn.* **2004**, *37*, 827–849.
6. Jason, G. Structured-light 3D surface imaging: a tutorial. *Adv. Opt. Photon.* **2011**, *3*, 128–160.
7. Je, C.; Lee, S.W.; Park,R.H. High-contrast color-stripe pattern for rapid structured-light range imaging. In *proc. 8th European Conf. Comp. Vision* **2004**, 95–107.
8. Maruyama, M.; Abe, S. Range sensing by projecting multiple slits with random cuts. *IEEE Trans. Pattern Anal. Mach. Intell.* **1993**, *15*, 647–651.
9. Guan, C.; Hassebrook, L.; Lau, D. Composite structured light pattern for three-dimensional video. *Opt. Express* **2003**, *11*, 406–417.
10. Ulusoy, A.O.; and Calakli, F.; Taubin, G. Robust one-shot 3D scanning using loopy belief propagation. *Computer Vision and Pattern Recognition Workshops* **2010**, 15–22.
11. Albitar, I.; Graebling, P.; Doignon, C. Robust structured light coding for 3d reconstruction. In *IEEE 11th Int. Conf. on Comput. Vision* **2007**, 1–6.
12. Koninckx, T.P.; Van Gool, L. Real-time range acquisition by adaptive structured light. *IEEE Trans. Pattern Anal. Mach. Intell.* **2006**, *28*, 432–445.
13. Zhang, L.; Curless, B.; Seitz,S.M. Rapid shape acquisition using color structured light and multi-pass dynamic programming. In *Proc. of 3DVT* **2002**, 24–36.
14. Benveniste, R.; Unsalan, C. Single stripe projection based range scanning of shiny objects under ambient light. In *proc. 24th Int. Symposium on Comput. and Inform. Sciences* **2004**, 1–6.
15. Rubner, Y.; Tomasi, C.; Guibas, L.J.; The Earth Mover's Distance as a Metric for Image Retrieval. *Int. J. Comput. Vision* **2000**, *40*, 99–121.
16. Fechteler, P.; Eisert, P. Adaptive colour classification for structured light systems. *IET Comput. Vis.* **2009**, *3* 49–59.
17. Cox, I.J.; Hingorani, S.L.; Rao,S.B.; Maggs, B.M. A maximum likelihood stereo algorithm. *Comput. Vis. Image Und.* **1996**, *63* 542–567.
18. Caspi,D.; Kiryati, N.; Shamir, J. Range imaging with adaptive color structured light. *IEEE Trans. Pattern Anal. Mach. Intell.* **1998**, *20* 470–480.
19. Sezgin, M.; Sankur, B. Survey over image thresholding techniques and quantitative performance evaluation. *J. Electon. Imaging* **2004**, *13* 146–165.
20. RB., Fisher; Naidu, DK. A comparison of algorithms for subpixel peak detection. In *Proc. of Advances in Image Processing, Multimedia and Machine Vision* **1996**, 385.
21. Wei, Z.; Jiang, S.; Ji, X. Survey reconstruction based on structured light of composite colorful stripes. In *Proc. of Conf. ICACTE* **2010**, *6* 123.

© July 21, 2012 by the authors; submitted to *Sensors* for possible open access publication under the terms and conditions of the Creative Commons Attribution license http://creativecommons.org/licenses/by/3.0/.