A new Robust Method of Line Detection in a Structured Light System

Hussam Yousef, Regis Huez, Laurent Hussenet and Michel Herbin
CReSTIC
University of Reims Champagne Ardenne
Reims, France
hussam.yousef@etudiant.univ-reims.fr, {regis.huez, laurent.hussenet, michel.herbin} @univ-reims.fr

***Abstract*— This paper presents an integrated 3D face scanning system using the structured light technique. A new pattern consisting of horizontal colored strips using the De_Bruijn sequence is designed in a manner that can ensure a minimal distance between any two strips of the same color. After illuminating the scene with this pattern, a first image is taken and used to obtain 3D information. To detect the strip centers in the captured image we use a smoothing Gaussian filter with a large kernel applied to the V component in the HSV color space. The size of this kernel is computed separately for each column. Then, a connection algorithm is used to connect the isolated points in the detected strips. The next step is to determine the colors of these detected strips. For this purpose we exploit the fact that the resulting strips are connected. Firstly, we use the H component of the HSV color space to determine the color of the easy sets of pixels in these strips. Then, we apply a region growing algorithm to assign the colors to the remaining pixels. Finally, each connected and colored part of strip is matched to a strip in the projected pattern and triangulated with its corresponding one in the projected pattern to generate the 3D model. Using the concept of connected strips we exploit the information from several columns to take the decision in each column, which enhances the robustness of the method used here versus the existed ones. A second image of the scene without illumination is obtained and used to add the texture to the reconstructed 3D model. Experiment results show an accurate 3D resolution with this technique.**

Keywords – Image Processing; Computer vision; 3D visualization; Structured light system.

I. INTRODUCTION

Recently 3D models are widely used. One of the first techniques in this field was the stereo vision. In this technique several images of the scene are taken from different points of view using a calibrated set of cameras. Each point of the scene is searched in the different images. Finally, triangulation step is applied to the detected points in the image to calculate their 3D position. The main drawback of stereo vision is the necessity of having landmarks [2].

Structured light systems are based on the association of a projector illuminating a scene with a coded pattern and a camera is used to capture one image of the illuminated scene. Each point of the coded pattern is determined in the captured image, and a triangulation step allows the attribution of the 3D positions. The first structured light systems used laser planes or laser dots scanning the object to be triangulated [7, 8]. These laser-based techniques allow a high resolution; nevertheless, mechanical operations will be required each time the laser changes its position.

Recently, the techniques based on a laser were replaced by a projector used to project a coded pattern. The image of the illuminated scene is captured and a 3D model is generated [9]. This technique can be separated into two basic groups [5]. In the first one, which is called time-multiplexing, a sequence of black and white patterns is projected on the scene; the drawback of this technique is that only static objects can be measured. The second one called one-shot technique uses only one colored pattern, and the unique position of each point is determined using its local neighborhood. In this context, it was shown [5] that a pattern using the De_Bruijn sequence achieves good results in terms of accuracy and resolution. A De_Bruijn sequence allows attributing a specific color to each line.

The generation of a coded color pattern using a De_Bruijn sequence can be achieved in two ways [2], with or without black gaps introduced between the colored lines (Fig. 1). The colored line projection generates colored stripes on the illuminated scene, therefore on the captured image. Both techniques require the detection of the strip centers to obtain a good 3D resolution after a triangulation step [7, 8, 9]. Using black gaps between strips gives the pattern more robustness against the variation of the ambient illumination.

In this paper, a structured light system using black gaps and applied to 3D face modeling is proposed (Fig. 1).

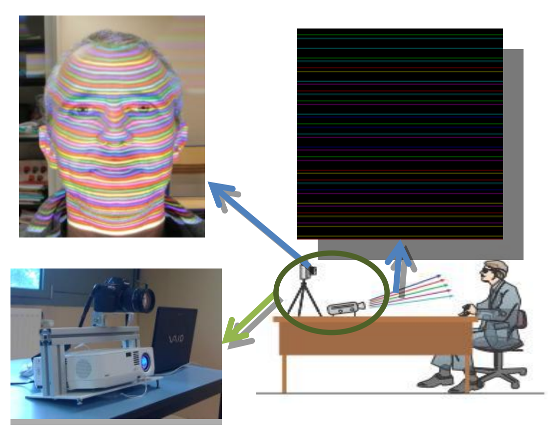

Figure 1. Layout of the system.

The part "Framework" will be firstly described, with a description of the used material. The second part named "location strip centers" concerns the determination of the center of the projected strips, this point is fundamental to obtain a sufficiently accurate 3D reconstruction. The following part talk about the line construction, where a global approach is taken contrary to a point to point approach. The last part "3D face model" presents the obtained results. A conclusion ends this paper.

II. FRAMEWORK

A. Overview

In a pattern consisting of horizontal colored lines, each column of the image can be seen as a 1D signal, its characteristics are: flat with bumps representing the projected lines. The determination of the projected lines requires information about the exact position of the bump centers. It also requires information on the colors of the determined centers, and finally a method to triangulate the centers with their corresponding points in the projected pattern.

The most popular methods to determine the pattern strips in structured light systems are based on the same principle. It consists in estimating on each column the possible candidates to be the centers of these lines and assigning them three probabilities:

- To be valid,
- To have a projected color,
- To be in a defined position in the used pattern.

The assignment of these points to specific points in the projected pattern [1, 2, 3] is carried out by methods of minimization of an objective function by dynamic programming. In this method, each column is treated separately without any consideration of the adjacent columns. It is obvious that the information from these columns could enhance the quality of the decision.

B. Method

We propose a global method to detect the centers of the strips and assign them to their corresponding points in the projected pattern. This method can be decomposed in four successive steps. In the first step, we determine the best position of the centers using a robust method, contrary to classic methods where a probability is given for all points. Secondly, the points are connected within their local neighborhoods. This procedure yields the entirely connected lines or at least long connected parts of the lines. The third and fourth parts are successively the color and lines assignments. In the last two steps, the detected strips are treated as an integrated unit instead of treating each one of its points alone.

C. Material

Fig. 1 shows the structured light system used, it consists of a « NEC NP410 LCD projector » with a full resolution of 1200*1600. It projects the colored multi-slit colored pattern on the face of the person at a distance of 0,4m. Then, the image of the illuminated scene is captured using a « Canon EOS 5D camera » with a resolution of 4368 by 2912 pixels. A polarized filter located before the camera can limit the problems of optical reflection and allows measurement taken in ambient light (Fig. 1). This system is controlled by a classic computer.

An offline calibration procedure is applied to synchronize the system and calculate the necessary projection matrices. The calibration procedure uses a dozen images of projected checkerboard on a plan containing a printed one. After the extraction of the corners in the projected and printed checkerboards, Zhang's method [10] is used to minimize the least square function over the needed parameters.

III. LOCATING STRIP CENTERS

A. Pattern generating

The patterns used in 3D reconstruction systems must be able to provide easy assignment of each point in the captured image to a point in the projected pattern. A minimal distance between any two strips having the same color must be kept to avoid unexpected connection of these similar strips in the captured image.

This application is dedicated to face reconstruction. Usually, the height of a face size is about 0,3m, so a pattern of 64 horizontal lines leads to a resolution of 4,6mm which is coherent with a face reconstruction. The 64-line pattern is achieved by a De_Bruijn binary sequence with a subsequence of six elements. This one is the starting point to construct a vector of 64 elements based on 6 different colors where the order of the colors is never repetitive. The colors are chosen to maximize the RGB color space. Two lines having the same color are separated by "at least" two lines of different color. The captured image is exploited in a RGB form.

B. Problem

The most famous way to locate strip centers is to search the peaks of strips in each separate column of each component RGB by fitting a 1D Gaussian profile in the neighborhood of the local maxima and searching its peak using the sub-pixel resolution [1, 2, 3]. With this method, two problems arise:

- In a color image, there is always an offset between the sub-pixel locations of the peaks in each RGB component, this problem is well known and is called RGB component misalignment [2, 4].
- The strips do not always have the same width, so two closed strips can be fitted by a unique Gaussian. On the other hand, a wide strip can be fitted by two Gaussians. The choice of the size of the Gaussian profile width should be carefully studied in order to ensure the result of this method.

In order to obtain a robust and performing solution, a global approach based on filtering is proposed. First of all a face shape detection step is applied to the image with the pattern to determine the region of interest, then we use a global approach to determine the strip centers using the V

 ISBN: 978-1-61208-184-7

component in the HSV color space. The V component is defined as:

$$V=\max(r,g,b) \tag{1}$$

This component has the biggest value of the three components in the RGB color space. In this way, we can get the center of the strip without considering color channels. Due to the RGB misalignment component, the V component signal can have some disturbed values at the strip centers.

In order to locate the strip centers, we use a smoothing filter with a big kernel; this filter smoothes the signal and focuses its peak in the middle. A common and powerful way to smooth a 1D signal is to convolute the signal with the Gaussian kernel.

Notice that a performing smoothing filter is similar to a low-pass filter with a low cut off frequency; it requires a big size of kernel, so a high value of the standard deviation. The size of the kernel k_s is defined as follows: the number of the nonzero values of the Gaussian kernel, the kernel is normalized and a truncation is used for all values lower to 0,01, k_s is proportional to the standard deviation. The influence of applying the Gaussian filter with different kernel sizes is depicted in Fig. 2. (Top) shows the intensities of the V component over a selected column, (middle) shows the same column after applying a Gaussian filter with a kernel of size 11, (below) shows the original column smoothed by a 21kernel size Gaussian filter.

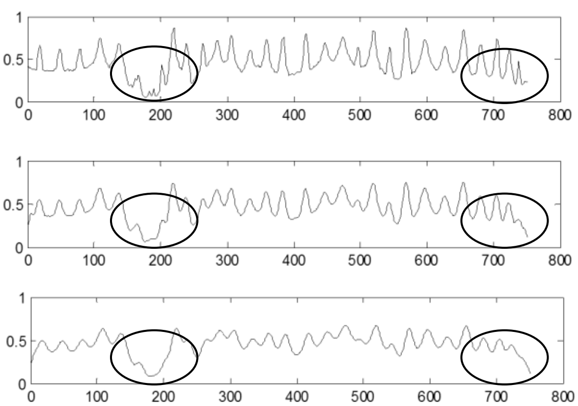

Figure 2. Application of kernels (size 11 and 21) on V component.

The larger the used kernel size the smoother the resulting signal, but a special attention should be paid to not confusing two successive strips and losing the relevant information as shown in Fig. 2 (below) when applying the kernel of size 21. To avoid this information lost, the size of the kernel must always be adjusted and controlled by the gaps between the strips in each scanned column.

Using this smoothing filter with a large kernel has two major advantages. The first one is that no small peaks correspond to the noise or disturbed strips could be found, while the second advantage is that this filter focuses the peaks of each strip exactly in the middle.

C. *Determination of the Kernel size k_s*

1) Algorithm

Due to the non-linearity of the scanned faces, the width of the gaps between the strips in the captured image is not always the same along the whole image. An important difference in size of these gaps can be found between the different columns. To yield the best results, a different kernel must be defined at each column according to the average of the gap sizes in this column.

Our contribution in this field is an algorithm dedicated to the determination of the kernel size for each separated column. On each column, the values of the pixels in the rising slope are summed up. The resulting values are associated to the positions of the relative peaks in the column. These values are normalized using the maximum of each column. A threshold is used so as not to take into account the low level values that are not representative of the strips. Fig. 3 describes two columns of the V component (dashed line), and the relative positions of the approximate peaks detected using this algorithm (solid line). The distances between each two successive peaks are calculated. To avoid the extreme abnormal values, the mean value of these distances is calculated and set as the size of the 1D kernel of the Gaussian filter that will be applied to the column.

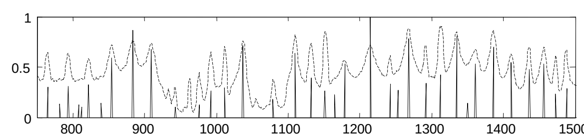

Figure 3. Approximate positions of the peaks in a column of V component.

The results present a variability of the gap between the peaks coming from several parameters, like the projector-subject distance or the surface inclination.

2) Algorithm validity

To test the validity of the algorithm, for several columns, the k_s value obtained by the algorithm is compared to the range of k_s values; which can give an exact restitution of the number of strips.

Over several images, a set of columns (10% of all) have been chosen to evaluate the validity of the proposed solution. This gives, for each image, a set of 71 columns covering all the areas of interest across the face.

The numbers of the projected colored lines on the face over these columns are counted using the human eye. Then for each of these columns we have applied a Gaussian filter with kernel size varying between 1 and 100, and compared the resulted peak number (found using the peak locating algorithm presented later) with the real number of lines. At the end of this step the range of k_s that yields the exact number of peaks is found for each one of the columns.

Fig. 4 shows an example for the image used, where the x-axis represents the index for the columns, the y-axis represents the values of k_s, the dashed and solid curves represent successively the minimum and the maximum of k_s that yields the exact number. The dotted curve shows the value of k_s found using the algorithm, and its related position to the range that yields the good results.

After this step of the determination of k_s, it is necessary to apply the Gaussian filter to the images.

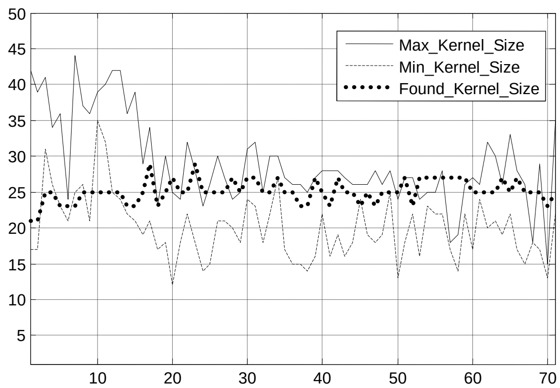

Figure 4. Ranges of values of k_s obtained on 71 columns

D. Locating peaks

1) Smoothing filter application

The Gaussian filter designed uses the previously described algorithm on each column of the image. As early mentioned this procedure smooths the signal and focuses the peaks of the strips in the middle. This procedure enables to recover the strip centers easily and with robustness. And it also makes the RGB component misalignment problem trivial. Fig. 5 shows several strips of a column signal, the solid curve represents the original signal while the dashed one represents the filtered signal.

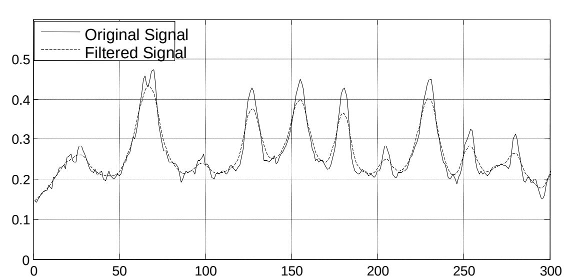

Figure 5. Original and filtered signal of a column signal.

On easy strips, the procedure does not improve the quality of the signal, in order to detect the local maxima, but on a bad quality signal (misalignment colors, etc.), this procedure allows to find a satisfactory solution.

2) Detection of the maximum

To determine the strip centers, it is necessary to find the local maxima along the smoothed signal of each column. The exact position of these local maxima is determined by applying the following simple algorithm to the smoothed signal; the difference between a point and the previous one must be positive, and the difference between a point and the following one must be negative, therefore the point is at the maximum of the curve.

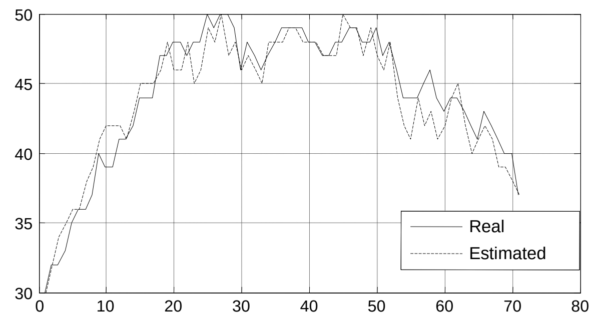

Figure 6. Peak detection error obtained on 71 columns.

In the special case where several points have the same value, the middle position (x-axis) is taken as maximum. This procedure is applied to all columns as previously. An evaluation of the detection quality is achieved by using the difference between the number of detected stripes and the real number of lines counted using the human eye. Fig. 6 shows the numbers of detected stripes, the solid and dashed curves are respectively the number of stripes detected by the eyes (real) and estimated by the algorithm. On several real images, the maximum of detection error is about 3 stripes. The mean of the error detection, number of real strips minus number of estimated strips, is about 1,5 strip/column.

IV. LINE CONSTRUCTION

A. From peaks to line

An image is constructed where each strip is represented by a simple line. Fig. 7 depicts the lines detected on a face; the one on the left represents the detected strips for the entire face superposed with the V component image. This image shows the exact position of the detected peaks in the middle of the strips. While the center and right parts of Fig. 7 successively represent a detail of the mouth and a detail of the nose. It is easy to understand that parts like the mouth will be easy to recognize and parts like the nose will be difficult to attribute.

 ISBN: 978-1-61208-184-7

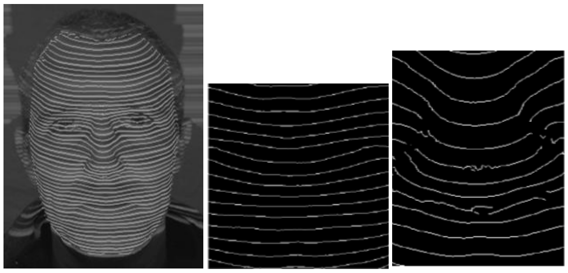

Figure 7. Line construction (face, mouth and nose).

B. Line connection

In the classic methods of structured light systems, the points resulting from the previous step are treated without connection. In order to increase the quality of the color attribution step and pattern affectation, a step of line connection is implemented to connect the isolated detected peaks within their local neighborhood. The horizontal projected lines in the captured image becoming curves have more or less accentuated slopes. As a result of searching the strip centers per column, a strip with a vertical slope generates discontinuity problems. Fig. 8 (top) shows the maximum allowed slope before getting this discontinuity. To overcome this problem the following connection algorithm is applied:

The points in the image are scanned from left to right and a criterion is tested in each one to decide whether it is well connected to the adjacent points in the line. If this point is found as an end of a connected line, a searching procedure is applied in a 7*10 pixel window to find the nearest beginning of another connected part of line. The two points are connected and the algorithm keeps scanning the other points in the image. An example of this problem and a resulting part of the image are shown in Fig. 8 (below).

The neighborhood used is determined empirically and found to be sufficient and yields good results in the face reconstruction application. At the end of this step, the majority of the detected peaks are connecting in long enough lines.

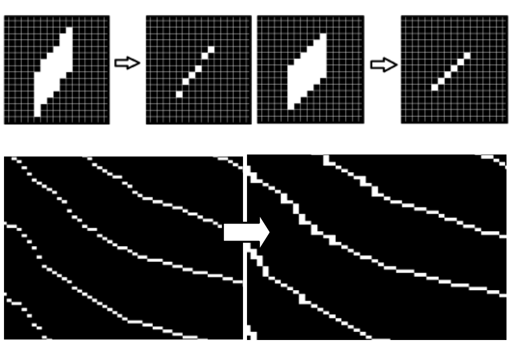

Figure 8. Top: maximum allowed slope, Down: connection algorithm.

C. Color of the line detection

1) Overview

It is well known that colors encounter a lot of distortions, starting with the influence of the projector, during its reflection on the surface of the measured object, and finally in the sensors of the camera. These distortions may be affected by either the absorption and reflecting proprieties of the measured object or by the proprieties of the projector and camera used which can lead to a cross-talk between the projector and the camera sensors. A lot of work has been done in this field; Zhang, Curless and Seitz [4] assumed that the surface is spectrally uniform and only calculated the projector-camera color cross-talk matrix, while in [3] an adaptive color classification is proposed to detect the colors of the projected lines using the RGB color space.

Several difficulties appear while using the RGB color space: the RGB components are highly correlated, the RGB space is not perceptually uniform, and it is highly sensible to color variations [6]. In our application, we chose to detect the color of the detected lines using the HSV color space and more precisely its H component which corresponds to the hue and refer to the dominant color. This color space is more related to the way in which human beings perceive colors and it is usually less affected by color variations.

In RGB color space, the colors of the pattern are chosen among the eight full saturated colors.

2) The H component histogram

Using the information of the H component and the connecting lines resulting from the previous step, the algorithm we used to assign the colors to these lines is described as the following:

The original captured image is converted to HSV color space and the H component is used. The ideal histogram of the six colors used in the H component is shown on Fig. 9a.

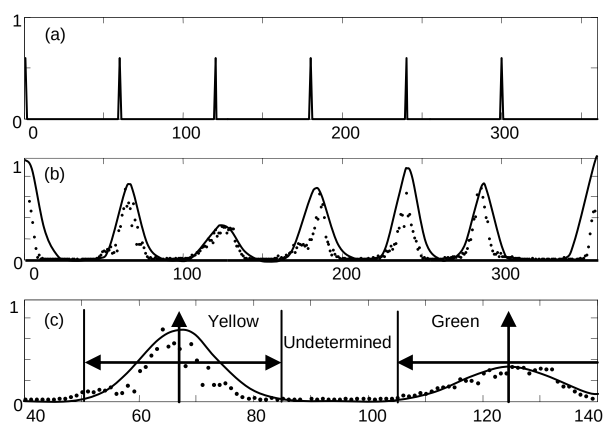

Figure 9. Histogram theoretical (a), experimental (b), enlargement (c).

The x-axis is graduated in degrees, so two successive colors are separated by 60°. The histogram of the detected lines in the H component image is shown in Fig. 9b. Due to

 ISBN: 978-1-61208-184-7

the distortion mentioned earlier, the distribution of each color is less regular than the ideal one. In most cases, the shape of the distribution is similar to a normal distribution around the mean value of each color area. The parameters of this distribution depend on the behavior of the colors on the skin, in other words, the interaction of light with skin. The modified parameters are the mean value, the shape and the width of the bumps. The bumps can be assimilated to Gaussian curves, so the width can be replaced by a standard deviation.

Due to these distortions, especially when these shapes have a small peak and large standard deviation, it is difficult to define a rule that can ensure a safe division of the H component into six regions and determine the boundaries between these regions.

The solution of this problem is divided into two steps: the first step is to find the pixels that can ensure a good color assignment, while the second one is to use the region growing in the connected lines to assign the colors of the remaining pixels.

3) *Color attribution*

The first step is achieved by dividing the distance between each two successive mean value positions in each color area into three equal regions.

Each one of the border regions has a safe color assignment to the related color; while the central region is labeled as undetermined (Fig. 9c) and set to the color white. In the second step, the indetermination will be lifted using the connected lines and the region growing.

The regions with the undetermined color are searched inside a connected line, and its color is found using the color of its adjacent pixels on both sides. These regions could be found in three possible cases:

- The regions can be within the connected line and surrounded by the colored regions from both sides and the surrounding regions have the same color on both sides, the undetermined color of the region is set as the same color.
- The regions can be within the connected line but the surrounding regions do not have the same color. This case is related to the connection of two lines with different colors. The regions with the undetermined color are deleted to avoid an error.
- Undetermined regions can be at the beginning or the end of the line, and then the color is determined by the color of the pixels on one side.

All the points in the connected lines resulting from this step have the same color. An example of the resulting image is shown in Fig. 10.

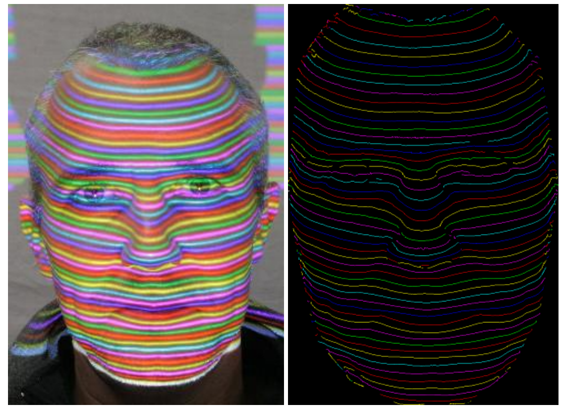

Figure 10. The resulting colored lines.

V. LINE MATCHING

As mentioned in the introduction of the paper, a De_Bruijn sequence is a cyclic sequence which consists of a certain number of subsequences. Each of these subsequences appears only once. This concept is used to match the detected strips to their corresponding ones in the projected pattern. The real scenes are not usually uniform ones. In many cases several parts of the projected lines will not be found in the captured image because of mislabeling, occlusions, shadows and other properties of each scanned object [4]. In this case, several colored lines will be missed in some columns. This means that the line matching using the classic dynamic programming is not always the best way. Multi-pass Dynamic Programming is proposed in [4]. In this method, the authors compute the monotonic components of the optimal path in multiple passes. To solve this problem, as mentioned earlier, each group of connected points is treated together as a connected line.

The information from all the points along the line is used to match this line to a line in the projected pattern. A subsequence of the six adjacent lines is formed and compared to De_Bruijn sequence used to create the projected pattern. In the case where some parts of the line have the problems mentioned earlier, such as occlusion, shadows and other problems will be matched using the information of the remaining points in the lines. For each of these lines the previous and following lines were found and used to assign it to a line in the projected pattern. The resulting points were triangulated with the corresponding points in the projected pattern by calculating the intersection of the two lines of sight defined by these points and the focal points of the camera and projector.

VI. 3D FACE MODEL

Several experiments have been performed. Fig. 11 shows typical example of these experiments; the first image with the pattern is used to generate the 3D wireframe model, while a second one without the pattern is used to add the texture.

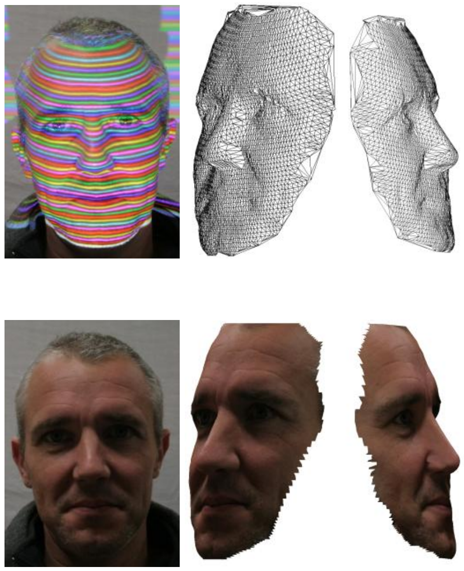

Figure 11. Images used and resulting wireframe model and textured one.

The step of peak detection yields some points which correspond to false peaks between the real peaks. Other eventual errors arise where the false peaks are found in the same local neighborhood. They can be connected during the connection step. In the step of color detection, a specific color will be assigned to these resulting false lines (the color red in most cases).

In the step of lines matching, the subsequence will be formed by this false line and the other five surrounded colored lines. As a result of use of the De_Bruijn sequence with big period (subsequence of six elements), there is no chance that the resulting subsequence will be matched to a subsequence in the De_Bruijn sequence, and these false lines or points will be excluded.

VII. CONCLUSION

We have presented in this paper our 3D scanner using the structured light system technique. Our robust method is applied on a human face; it can be also used in many other fields, either in computer vision or industry applications. A specially designed color pattern using the De_Bruijn sequence is proposed to enhance the robustness of the system.

A new method to locate the strip centers in the captured image uses a smoothing filter with a large kernel applied to the V component of the HSV color space. The validation step of the algorithm used to choose the size of the kernel has been presented, and satisfactory results have been obtained. Usually, each strip center is treated alone, and more precisely, each point of the picture is treated alone. We use an innovative and global technique in order to construct usable lines. The method is based on a growing algorithm using an appropriate window and it is applied to V component image. We treat each set of connected points together while determining their color and assigning them to a line in the projected pattern. The last paragraph presents the results obtained on 3D face modeling. A texture has been applied to be closer to reality.

Due to the robustness of the proposed 3D face acquisition method it has been used to scan faces for web applications.

The week point of this work is the absence of comparative studies with the existing approaches. Such studies form the major part of our future work.

REFERENCES

[1] P. Fechteler, P. Eisert, and J. Rurainsky, "Fast and High Resolution 3D Face Scanning" In Proc. of the International Conference on Image Processing (ICIP 2007), vol. 3, pp. 81-84, San Antonio, Texas, USA, doi:10.1109/ICIP.2007.4379251.

[2] J. Pages, J. Salvi, C. Collewet, and J. Forest, "Optimised De Bruijn patterns for one-shot shape acquisition" Image and Vision Computing 2005, vol. 23, Issue 8, pp. 827-849, doi:10.1016/j.imavis.2005.05.007.

[3] P. Fechteler, P. Eisert, "Adaptive Color Classification for Structured Light Systems", in Computer Vision and Pattern Recognition Workshops (CVPRW 2008), pp. 1-7, doi:10.1109/CVPRW.2008.4563048.

[4] L. Zhang, B. Curless, and S.M. Seitz, "Rapid shape acquisition using color structured light and multipass dynamic programming", 1st IEEE International Symposium on 3D Data Processing 2002, Visualization, and Transmission, pp. 24–36, Padova, Italy, doi:10.1109/TDPVT.2002.1024035.

[5] J. Salvi, J. Pagès, and J. Batlle, "Pattern codification strategies in structured light systems", Pattern Recognition 2004, vol. 37, Issue 4, pp. 827-849, doi:10.1016/j.patcog.2003.10.002.

[6] Jiebo Luo, D. Crandall, "Color object detection using spatial–color joint probability functions", IEEE Transactions on Image Processing 2006, vol. 15, Issue 6, pp. 1443-1453, doi:10.1109/TIP.2006.871081.

[7] D. Scharstein, R. Szeliski, R. Zabih, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms", In Proc. of IEEE Workshop on Stereo and Multi-Baseline Vision, SMBV 2001, pp. 131-140, doi:10.1023/A:1014573219977.

[8] A. Dipanda, S. Woo, "Towards a real-time 3D shape reconstruction using a structured light system", Pattern Recognition, Oct. 2005, vol. 38, Issue 10, pp. 1632-1650, doi:10.1016/j.patcog.2005.01.006.

[9] Y. F. Li and B. Zhang, "A Method for 3D measurement and reconstruction for active vision", Measurement Science and Technology, 2004, vol. 15, Issue 11, pp. 2224-2232, doi:10.1088/0957-0233/15/11/007.

[10] Z. Zhang, "A flexible new technique for camera calibration", IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, vol. 22, Issue 11, pp. 1330-1334, doi:10.1234/12345678.