Face Tracking Based on Skin Color and Motion Prediction

Jing Sun

College of Information and Electronics
Zhejiang Sci-Tech University
Hangzhou, China
jings531@163.com

Abstract—**Aimed at the disturbance caused by illumination variations and similar skin color regions in the video scene which will greatly affect face tracking accuracy, and in order to enhance the real time performance, a novel approach for face tracking based on skin-color information and motion prediction is put forward in this paper. The major idea is to take the Gaussian skin-color model based on YCrCb space as face location model, and in order to increase the real-time performance of skin-color detection, GM(1,1) model is combined to carry out face prediction so as to reduce searching areas. By exploiting least-square method to determine constant c in time response function of the GM(1,1) whitenization function, the prediction accuracy can be greatly advanced, satisfying the demand for prediction model owing to the randomization of human face movement. Experimental results show that the proposed method has such advantages as fast tracking speed and good robustness.**

Keywords-Skin-color Information, Motion Prediction, Face Tracking, GM(1,1) Model

I. INTRODUCTION

With the rapid development of information processing technology and the need of various applications such as video surveillance, face tracking [1-4] has become a major concern in the field of computer vision. It has broad prospects in various areas such as video surveillance, human–computer interface, face recognition, pose estimation, facial expression recognition, content-based image retrieval and image coding, etc. Face tracking using skin-color information as tracking cue is popular with many researchers for its rapid tracking speed and good robustness, as well as its independence of facial features and not affected by the change of facial pose, scale and expression. However, in practice, a single skin-color model can hardly confront the challenge of nature circumstances. This is because face tracking can easily be influenced by many factors such as similar skin-color regions, non face-skin areas, light variations in background also with human movement and dissimilar performance of camera equipment and so on. To complement this defect and meanwhile to accurately detect the face position in real-time, a new method for face tracking based on skin-color information and an improved grey prediction model is presented in this paper.

II. SKIN COLOR MODEL BASED ON YCRCB SPACE

Skin color is one of the most critical features of human face. To utilize the skin color features, firstly, the image is transformed into YCrCb space, which has the similar constructional principle with the process of human visual perception, and its major benefits come from the ability to separate luminance from chrominance as HSI Space can do, as well as its simple calculation process and spatial coordinate representation, all this leading to best color clustering ability. Through the nonlinear transformation in YCrCb space, making the skin cluster independent of illumination component, we can get ideal color detection result. The conversion formula is as follows [5]:

$$\overline{C_b}(Y) = \begin{cases} 108 + \dfrac{(K_l - Y)\cdot(118-108)}{K_l - Y_{\min}} & if(Y < K_l) \\ 108 + \dfrac{(Y - K_h)\cdot(118-108)}{Y_{\max} - K_h} & if(K_h < Y) \end{cases} \quad (1)$$

$$\overline{C_r}(Y) = \begin{cases} 154 - \dfrac{(K_l - Y)\cdot(154-144)}{K_l - Y_{\min}} & if(Y < K_l) \\ 154 + \dfrac{(Y - K_h)\cdot(154-132)}{Y_{\max} - K_h} & if(K_h < Y) \end{cases} \quad (2)$$

$$Wc_i(Y) = \begin{cases} WLc_i + \dfrac{(Y - Y_{\min})\cdot(Wc_i - WLc_i)}{K_l - Y_{\min}} & if(Y < K_l) \\ WHc_i + \dfrac{(Y_{\max} - Y)\cdot(Wc_i - WHc_i)}{Y_{\max} - K_h} & if(K_h < Y) \end{cases} \quad (3)$$

$$C_i'(Y) = \begin{cases} (C_i(Y) - \overline{C_i(Y)})\cdot\dfrac{Wc_i}{Wc_i(Y)} + \overline{C_i(Y)} & if(Y < K_l) or(K_h < Y) \\ C_i(Y) & if(Y \in [K_l, K_h]) \end{cases} \quad (4)$$

where $K_l = 125$, $K_h = 188$, $Y_{\min} = 16$, $Y_{\max} = 235$, i indicates b or r, and $Wc_b = 46.97, WLc_b = 23, WHc_b = 14$, $Wc_r = 38.76$, $WLc_r = 20, WHc_r = 10$. So, after the piecewise non-linear color transformation we can use an ellipse to depict the skin-color regions approximately .The formula of skin-color model is:

$$\frac{(x - ec_x)^2}{a^2} + \frac{(y - ec_y)^2}{b^2} = 1 \quad (5)$$

where $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} c_b - c_x \\ c_r - c_y \end{bmatrix}$ (6)

the coefficient constant $ec_x, ec_y, c_x, c_y, \theta, a$ and b are 1.60, 2.41, 109.38, 152.02, 2.53(radian), 25.39 and 14.03 respectively.

The image after color segmentation, however, inevitably contains some other small areas besides the face regions. In order to reject these disturbance areas, the image can be regarded as being composed of several connected regions, the 8-connected regions can be marked up and whose pixel number exceeding a certain value will be determined as the face area. In the mean time, to deal with the situation such as light changing (gradual change, sudden change, reflective and so on) in complex environment, also with the happening of false and wrong detected faces due to the shadow caused by object movement and the impact of similar skin color on face tracking, as well as to improve the real time performance of face tracking by avoiding the global search of skin-color model, we introduce the GM(1,1) model to make motion prediction, so as to get a robust tracking result.

III. GREY PREDICTION MODEL GM(1,1)

GM(1,1) is a popular grey model for dynamic prediction consisting of a single variable and a first-order differential equation. Suppose $X^{(0)}$ to be a non-negative incremental original sequence with equal time interval as follows:

$$X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)) \qquad (7)$$

Where n is the length of the sequence. To strengthen the regularity and weaken the randomness, the initial sequence must be transformed. The most popular way is the first-order accumulated generating operation, an approach making grey process from grey to white, that is to say, from which the developing trend of the accumulation process of grey variables can be found, revealing the interior laws implied in the original chaotic data. The first-order accumulated sequence $X^{(1)}$ is

$$X^{(1)} = (x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)) \qquad (8)$$

where, $x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i)$, k=1,2,…,n; $\qquad (9)$

Then the grey differential equation

$$x^{(0)}(k) + az^{(1)}(k) = b \qquad (10)$$

can be called the original form of GM(1,1) model.

The background value $Z^{(1)}$ is defined as

$$Z^{(1)} = (z^{(1)}(1), z^{(1)}(2), \cdots, z^{(1)}(n)) \qquad (11)$$

Usually it is constructed by the trapezoidal form of $X^{(1)}$, i.e.

$$z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k-1) \qquad (12)$$

If $\hat{a} = (a,b)^T$ be a parameter vector, and

$$Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}, \quad B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix} \qquad (13)$$

then the least square estimation parameters of the GM(1,1) model $x^{(0)}(k) + az^{(1)}(k) = b$ satisfy the equation

$\hat{a} = (B^T B)^{-1} B^T Y$, and we get the explicit expression of estimation parameter a and b:

$$a = \frac{CD - (n-1)E}{(n-1)F - C^2}, \quad b = \frac{DF - CE}{(n-1)F - C^2} \qquad (14)$$

where

$$C = \sum_{k=2}^{n} z^{(1)}(k) \quad D = \sum_{k=2}^{n} x^{(0)}(k)$$

$$E = \sum_{k=2}^{n} z^{(1)}(k)x^{(0)}(k) \quad F = \sum_{k=2}^{n} z^{(1)}(k)^2 \qquad (15)$$

so, the time response formula of the whitenization equation $\dfrac{dx^{(1)}}{dt} + ax^{(1)} = b$ is

$$\hat{x}(t) = ce^{-at} + \frac{b}{a}, \quad t=1, 2, ...,n \qquad (16)$$

where

$$A = \begin{bmatrix} x^{(1)}(1) - \dfrac{b}{a} \\ x^{(1)}(2) - \dfrac{b}{a} \\ \vdots \\ x^{(1)}(n) - \dfrac{b}{a} \end{bmatrix} \quad D = \begin{bmatrix} e^{-a} \\ e^{-2a} \\ \vdots \\ e^{-na} \end{bmatrix}$$

$$c = (D^T D)^{-1} D^T A,$$

The time response formula of the grey differential equation in GM(1,1) model is

$$\hat{x}^{(1)}(k) = ce^{-ak} + \frac{b}{a}, \quad k=1,2,...,n \qquad (17)$$

And the prediction value is

$$\hat{x}^{(0)}(k+1) = \alpha^{(1)} \hat{x}^{(1)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) \qquad (18)$$

We know that the essential of GM(1,1) to make prediction is the process which first using the solved value $\hat{x}^{(1)}$ to fit $x^{(1)}$,then get $\hat{x}^{(0)}$ restored from $\hat{x}^{(1)}$ by regressive reduction operation. However, the solved prediction formula may not be the optimal one because it is shown that to force $x^{(0)}(1)$ as the initial value in the traditional GM(1,1) model lacks rigorous theoretical basis [7], for the fitting curve $\hat{x}^{(1)}$ acquired by this method will definitely pass through the point $(1, \hat{x}^{(1)}(1))$ in the coordinate plane $(k, \hat{x}^{(1)}(k))$. Whereas according to the least-square theory, the fitting curve could not pass the first point, hence the theoretical basis of taking $\hat{X}^{(1)} = X^{(1)}(1)$ as a known condition does not exist. On the other hand, we should take into account that $X^{(1)}(1)$ is the oldest datum neither having close relationship to the future nor generated by accumulated generating operation thus resulting weak regularity. The improvement of simply choosing other data as the initial value is not thorough. Reference [8] pointed out that using the first component of sequence $X^{(1)}$ as the initial condition of the grey differential model will cause insufficient utilization of the new information. In accordance with the idea incorporated in the

grey system theory that new information have larger priority, the performance of grey modeling can be improved by giving the new information larger weight in the process of modeling. So it proposed applying the nth-component $x^{(1)}(n)$ of $X^{(1)}$ instead of $x^{(0)}(1)$ as the initial condition of the grey differential model in the traditional GM(1,1) model.

But we also note that using $x^{(1)}(n)$ may not be the best choice because of stochastic error and bad data which is likely to occur in the processing. For the above reasons, [6] suggested that the optimization of time response function with certain component as initial condition can hardly ensure the best fitting between the initial and the simulation sequence. Meanwhile, these approaches not only make time response function highly rely on the developing coefficient but also cannot guarantee achieving the least simulation error of the whole sequence. To overcome the problem, after getting developing coefficient a and the grey component b, the constant c in the whitenization equation can be determined by minimizing the sum of squares of the difference between $x^{(1)}$ sequence and its simulation values. As a result, the optimal time-response function satisfying the initial sequence can be built, significantly improving the accuracy.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

Grey system is an uncertain system with the feature that "some information known while some unknown". The stability and accuracy of grey prediction model have been an exciting challenge for grey system theory. Staring from this point, ensuring the stability of grey prediction is the hypothesis for exploiting grey prediction model. Moreover, only with improved prediction accuracy and reduced prediction error, can the wider application of grey model exists. Consequently, by computing the constant c in the whitenization function of the GM(1,1) model using the least-square method, the prediction accuracy can be improved, meeting the need of prediction model respect to the randomization of human face movement, avoiding the invalidation of the prediction model caused by sudden change of facial motion direction and abrupt acceleration.

While face tracking involves determining relationships of faces in each frame of image sequences and achieving the information such as position, size and direction of the detected face in the following frames. When considering face tracking in complex, wide dynamic variation background, it is unreliable only using Gaussian skin-color model because the occurrence of similar skin-color region or other moving objects. Therefore, the initial searching point can be acquired by motion prediction, and the shorter the distance between which and the global extreme point, the smaller probability the searching algorithm disrupted by the local extreme point, greatly reducing the risk of disturbance due to the complex surrounding, as well as significantly decreasing calculation. In this paper, the on-line face tracking system is built combined with skin-color information and the GM(1,1) prediction model. The experimental result is shown in Fig 1, from which we can see it apparently that the face in the

image sequences is making obviously randomize movement, and under the approach proposed in our paper the face tracking can be implemented with great stability.

(a) Frame 10   (b) Frame 20
(c) Frame 25   (d) Frame 30
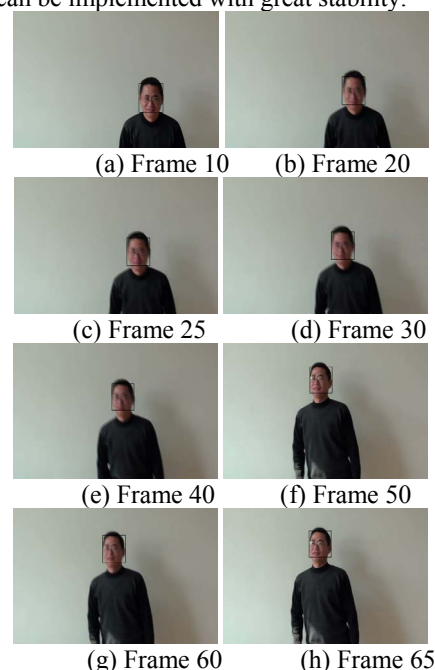(e) Frame 40   (f) Frame 50
(g) Frame 60   (h) Frame 65

Figure 1.   The result of face-tracking

In summary, the face tracking system combined skin-color model and motion prediction can track face of randomized movement in real-time with high accuracy. When the moving face obscured by other objects, the approach suggested here can track continuously by replacing the real value with the prediction value of the GM(1,1) model.

V. CONCLUSION

For the colorful face image in a complex environment, face region to be detected can be obtained through the process that first segmenting the skin color regions from the target image then making rearrangement based on the skin-color information. To deal with the light variation in the complex circumstance and similar skin color region in the background, a mechanism combined the idea of motion prediction and GM(1,1) to forecast the target location in next frame so as to decrease the searching area is presented in this paper, which leads to both the elimination of disturbance and increasing real time performance as well as good robustness in face tracking. Additionally, using least square method to get constant c of the time response function in the GM(1,1) whitenization function, ensuring the best fit between the initial and simulated sequence, the prediction accuracy can be improved substantially, greatly fitting the need for prediction model towards the randomization of human face movement.

REFERENCES

[1] WU Tian-rui; ZOU Yue-xian; WANG Wei. Skin-color based particle filtering for human face tracking. 2008 IEEE International Conference on Cybernetics and Intelligent Systems, 2008, pp:728-733

[2] WANG Chuan-xu; Li Zuo-yong. A new face tracking algorithm based on local binary pattern and skin color information. Proceedings - International Symposium on Computer Science and Computational Technology, 2008, v 2, pp: 657-660

[3] Essannouni L.; Elhaj E. Ibn; Aboutajdine D. Automatic color face tracking using gray information. Proceedings - IEEE Symposium on Computers and Communications, 2008, pp: 1010-1014

[4] ZHOU Zhi-yu; XIA Jin-song; WU Lang; ZHANG Jian-xin. Robust Face Tracking based on Mean Shift Algorithm and GM(1,1) Model. Journal of Computational Information Systems[J], 2010,6(1) , pp : 263-270

[5] Jain A K; Yu Z; Lakshmanan S. Object matching using deformable template. IEEE Transactions on Pattern Analysis and Machine Intelligence[J], 1996, v 18 ,n 3, pp:267-278

[6] LIU Bin; ZHAO Liang; ZHAI Zhen-Jie; et al. Optimum model of GM(1,1) and its suitable range. Journal of Nanjing University of Aeronautics and Astronautics[J], 2003, v 35, n 4, pp: 451-454

[7] ZHANG Da-hai; JIANG Shi-fang; SHI Kai-quan. Theoretical defect of grey prediction formula and its improvement. Systems Engineering theory & Practice[J], 2002, v 22, n 8, pp:140-142

[8] DANG Yao-guo; LIU Si-feng; LIU Bin. The GM Models that x(1)(n) be Taken as Initial Value. Chinese Journal of Management Science[J], 2005, v 13, n 1, pp:132-135