

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
23 June 2011 (23.06.2011)

PCT

(10) International Publication Number
WO 2011/074014 A2

(51) International Patent Classification: Not classified

(21) International Application Number:
PCT/IN2010/000823

(22) International Filing Date:
16 December 2010 (16.12.2010)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
2905/MUM/2009 16 December 2009 (16.12.2009) IN

(71) Applicant (for all designated States except US): **TATA CONSULTANCY SERVICES LTD.** [IN/IN]; Nirmal Building, 9th Floor, Nariman Point, Mumbai 400 021, Maharashtra (IN).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **BHOWMICK, Brojeshwar** [IN/IN]; Bengal Intelligent Park, Building D, Plot No. A2 M2 & N2, Block-EP, Salt Lake Electronics Complex, Sector V, Kolkata 700 091 West Bengal (IN). **CHIDANAND, K.S.** [IN/IN]; Bengal Intelligent Park, Building D, Plot No. A2 M2 & N2, Block-EP, Salt Lake Electronics Complex, Sector V, Kolkata 700 091 West Bengal (IN).

(74) Agent: **MOHAN, Dewan**; R.K. Dewan & Company, Trade Mark & Patent Attorneys, Podar Chambers, S.A. Brelvi Road, Fort Mumbai 400 001, Maharashtra (IN).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

- as to the identity of the inventor (Rule 4.17(i))
- of inventorship (Rule 4.17(iv))

Published:

- without international search report and to be republished upon receipt of that report (Rule 48.2(g))

(54) Title: A SYSTEM FOR LIP CORNER DETECTION USING VISION BASED APPROACH

(57) Abstract: A system and method for lip and lip corner detection using computer vision based techniques have been disclosed. The proposed system uses an integral image to extract Haar like features for face detection. Further, the system employs edge detection and clustering techniques to localize the lips on the detected face. A weighted sum of square difference technique is then applied to the lip region to determine the lip corners. The system does not depend on color models, contour extraction models or template matching techniques for lip and lip corner detection. Hence, the system is not sensitive to contour and color models and enables rapid and accurate lip and lip corner detection in real time.



WO 2011/074014 A2

A SYSTEM FOR LIP CORNER DETECTION USING VISION BASED APPROACH

FIELD OF THE INVENTION

The present invention relates to the field of Computer Vision and Image Processing.

Particularly, the present invention relates to a system and method for detecting lip corners for a makeup / model company based on computer vision technologies.

DEFINITIONS OF TERMS USED IN THE SPECIFICATION

The term 'cascaded classifiers' in this specification relates to concatenation of two or more classifiers, that is output of one classifier is given to the subsequent classifier.

The term 'Haar-like features' in this specification relates to features that indicate specific characteristics in an image and can be utilized to recognize objects like a facial region in an image.

The term 'histogram' in this specification relates to a graphical representation which plots the number of pixels in an image on the vertical axis with their brightness value on the horizontal axis.

The term 'horizontal dilation' in this specification relates to an expansion or transformation of an object horizontally about its x-axis based on a factor value.

The term 'integral image' in this specification relates to a representation of an image obtained by performing certain mathematical operations on the pixels of that image. Typically, a pixel in the representation is obtained as the sum of the current pixel value with all pixels above and to the left of it. The 'integral image' is also known as summed area tables.

The term 'sobel mask' in this specification relates to a mask applied to a gray scale image to detect edges in an image.

BACKGROUND OF THE INVENTION AND PRIOR ART

Detection of lips and lip-corners is gathering enormous attention in today's world especially in the cosmetic and medical domains.

Most of these lip detection techniques rely more on colour model, specifically HSV colour model. Although some of the approaches based on visual technology, these techniques use equations which detect lips only for a select few colour images and not for all colour images or gray scale images. Also, the existing technology is quite expensive with regards to the hardware need and the strong computational requisites for image processing and lip detection techniques.

Existing techniques for lip detection are the following:

No.	Title	Approach
1	Lip detection and tracking. By A. Caplier Published in IEEE	This publication discloses the use of active contour model to detect lips.

2	Lip detection by the use of neural networks. By Jamal Ahmad Dargham et.al Published: Springer Japan Issue: Volume 12, Numbers 1-2 / March, 2008	This publication discloses the use of neural network based training. The training utilizes supervised learning.
3	Lip Detection Using Confidence-Based Adaptive Thresholding. By Jin Young Kim Publisher Springer Volume 4291/2006	This publication discloses the transformation of images with respect to hue and then employs fuzzy rule based approach to take decision about the presence of lips.

There have also been attempts in the prior art to detect lip corners. For instance United States Patent Application 20030212552 discloses a method for feature extraction to accurately determine the lip position. The disclosed method identifies the facial image from an audio video stream using Gaussian mixture model to model the color distribution of the face region. The generated color distinguished face template along with a background region logarithmic search is used to deform a face template to fit on the face to optimally identify face(s) in a visual scene. From the identified face, the mouth region is segmented. Further, linear discriminant analysis is performed on the mouth region to identify the lips. The contour of the lips is determined through binary chain encoding and the mouth corner using a corner finding filter. However, the disclosure relies on a face template for face detection and active contour model to detect lips.

Similarly, United States Patent Application US2007154095 and United States Patent Application 20070154096 employ an AdaBoost algorithm combined with rectangular filters to detect corners of the mouth. However, even their disclosures depend upon the conventional contour extraction algorithm applied to a binary mask on a detected candidate mouth region to determine the mouth corners as the left most and the right most points in the contours of the binary mask.

There is therefore a need for a lip corner detection system which in real-time detects the corners of the lips without using the conventional template matching and contour extraction techniques. There is also a need for high speed and an efficient system which accurately identifies lip corners.

OBJECT OF THE INVENTION

An object of this invention is to provide a system and a method for computation of lip localization in real-time.

Another object of this invention is to provide an accurate system for detection of lip corners.

Yet another object of this invention is to provide an economical system for lip corner detection.

Still another object of this invention is to provide a system for lip corner detection which requires minimum hardware.

One more object of this invention is to provide a system for lip corner detection which is time efficient.

Further another object of the present invention is to provide a system for lip corner detection for use in make up model company applications.

SUMMARY OF THE INVENTION

The present invention envisages a system for lip and corner detection using a color or a gray scale image. The proposed system aims at detecting lip corners of a model / a user by employing computer vision techniques.

The system envisaged by the present invention for lip and lip corner detection comprises the following components:

- a face detection unit having at least one haar classifier adapted to receive and detect a face region in an image;
- a lip localization unit adapted to receive a detected face region and further adapted to locate a lip area;
- a clustering unit adapted to receive data related to the lip area from the lip localization unit and obtain a lip boundary; and
- a lip corner detection unit having processing means adapted to receive data related to the lip boundary and extract lip corners.

Typically, the face detection unit includes:

- intermediate image representation formation means adapted to receive the image in the pixel form and further adapted to apply predefined operations on pixels of the image and still further adapted to obtain an intermediate image representation for the image;
- feature extraction means having window creation means adapted to receive the intermediate image

representation and further adapted to create sub windows of predefined sizes over the image to extract predetermined 'Haar like features' from the sub windows and still further adapted to provide extracted features for each of the sub windows; and

- a plurality of cascaded haar classifiers arranged in predetermined format adapted to receive the extracted features and classify the sub windows as a positive image or a negative image, wherein a positive image indicates a face region and negative image indicates a non-face region.

Preferably, the intermediate image representation formation means is adapted to extract an intermediate representation of an image using the Integral Image technique.

In addition, the cascaded haar classifiers are selected from the group of classifiers consisting of:

- at least one simple classifier adapted to reject sub-windows containing the negative images and further adapted to forward the candidate positive images to a succeeding classifier; and
- at least one complex classifier adapted to receive and reject the candidate positive image in the event that the image is a false positive image.

Further, the cascaded haar classifiers include:

- a features repository adapted to store a set of predetermined features;

- a training repository adapted to store a set of positive and negative images, wherein a positive image indicates a face region and negative image contains a non-face region; and
- classification means co-operating with the features repository and the training repository adapted to receive the extracted features for each sub-window of the image and further adapted to reject the sub-windows containing the negative images and accept the sub-windows containing the positive images.

Additionally, the face detection unit further includes:

- histogram generation means adapted to receive the positive image and generate an operative vertical histogram and an operative second horizontal histogram, wherein the vertical histogram and the second horizontal histogram represent intensity values in operative horizontal and vertical directions;
- filter application means adapted to remove speckle noises from the vertical histogram and the second horizontal histogram; and
- computational means adapted to receive the vertical histogram and compute an average of the vertical intensity values and further adapted to mark peaks and valleys in the vertical histogram and still further adapted to detect the center of the face as the peak value in the vertical histogram of the positive image.

Still Further, the lip localization unit includes:

- binarization means adapted to convert the positive image into a gray scale image;
- masking means adapted to receive and apply an operative horizontal Sobel mask to the gray scale image and further adapted to provide a masked image;
- histogram generation means adapted to generate an operative first horizontal histogram for the masked image, wherein the first horizontal histogram represents pixel intensity values on an operative horizontal axis; and
- lip detection means adapted to receive the first horizontal histogram and identify peaks in the first horizontal histogram and further adapted to mark the peaks as the lip area.

Furthermore, the clustering unit includes:

- lip edge dilation means adapted to perform operative horizontal dilation on the lip area and further adapted to obtain a candidate lip region, wherein the candidate lip region includes candidate points representing the candidate lip region;
- clustering means adapted to apply a predetermined clustering technique on the candidate points and further adapted to obtain a lip boundary; and
- filtering means adapted to apply a predefined filter to points in the lip boundary and further adapted to obtain a filtered gray scale two dimensional lip region image.

Typically, the clustering means performs the clustering operation using k-means clustering technique.

Preferably, the processing means of the lip corner detection unit is adapted to perform a mathematical operation of 'weighted sum of square difference' over the lip boundary and still further adapted to perform non-maximal suppression technique to identify lip corners.

In accordance with the present invention there is provided a method for detecting lips and lip corners, the method comprising the following steps:

- detecting a face region of a user in an image;
- localizing a lip area in the image;
- receiving data related to the lip area and clustering the data to obtain a lip boundary; and
- mathematically processing data related to the lip boundary and extracting lip corners.

Typically, the step of detecting a face of a user in the image includes the following steps:

- forming an intermediate image representation for the image;
- creating sub windows of a predefined size over the image;
- extracting predetermined features for the image from each of the sub windows; and
- classifying the sub windows as positive image or negative image, wherein a positive image indicates a face region and negative image indicates a non-face region.

Preferably, the step of forming an intermediate image representation for the original image includes the steps of:

- selecting pixel coordinates in the original image, one at a time; and
- computing a value for each one of the corresponding pixel coordinates in the intermediate image representation as the sum of pixels located in the pixel coordinates above and to the left of a current pixel whose value is being computed.

In accordance with the present invention, the step of classifying the sub windows as positive image or negative image includes the step of verifying the sub-windows by passing them through a series of cascading haar classifiers.

Further, the step of localizing a lip area in the image includes the following steps:

- converting the positive image into a gray scale image;
- masking the gray scale image by applying an operative horizontal Sobel mask;
- generating an operative first horizontal histogram for the masked image, wherein the first horizontal histogram represents pixel intensity values on a horizontal axis; and
- identifying peaks in the histogram and marking the peak as the lip area.

Still further, the step of receiving data related to the lip area and clustering the data to obtain a lip boundary includes the steps of:

- performing an operative horizontal dilation on the lip area to obtain a candidate lip region, wherein the candidate lip region includes candidate points representing the candidate lip region;
- performing clustering operation on the candidate lip region to obtain a lip boundary; and
- applying a predefined filter to the points in the lip boundary to obtain a filtered gray scale lip boundary image.

Furthermore, the step of mathematically processing data related to the lip boundary and extracting lip corners includes the following steps:

- creating an original image patch of a predetermined size;
- shifting the patch by a predefined factor to obtain a new shifted patch; and
- computing the weighted sum of square difference between the original image patch and the new shifted patch to identify the similarity between the two patches.

BRIEF DESCRIPTION OF THE ACCOMPANYING DRAWINGS

The present invention will now be described with reference to the accompanying drawings, in which:

FIGURE 1 illustrates a schematic of the system for lip corner detection using vision based techniques in accordance with the present invention;

FIGURE 2 illustrates a schematic display of cascading classifiers in accordance with the present invention;

FIGURE 3 illustrates a 5×5 smoothing mask matrix in accordance with the present invention;

FIGURE 4 illustrates a flowchart showing the steps involved in lip and lip corner detection in accordance with the present invention; and

FIGURE 5 is a flowchart showing the detailed steps involved in lip and lip corner detection in accordance with the present invention.

DETAILED DESCRIPTION

The system for vision based approach for lip corner detection will now be described with reference to the accompanying drawings which do not limit the scope and ambit of the invention. The description provided is purely by way of example and illustration.

The existing techniques for lip and lip corner detection are based on template matching techniques firstly for detection of the face and then for localizing the region of the lips. However, templates are not very efficient as they involve searching in a template database for a best fit template for a face model. Thus, making the system not very time efficient and lowering the speed of real-time face detection. Also, after the face has been detected, contour extraction models / color models are employed for lip corner detection. Thus, to make the real-time processing faster and more accurate the present invention proposes a system and a method to identify lips and the lip-corners on a human face through the usage of machine/computer vision and image processing techniques.

The present invention provides an efficient system and method which in real time identifies the definitive lip shape and lip corners. One of the applications of the present invention includes detecting shape of the lips and the lip corners for a Makeup Model company. In this application, the

proposed system enables the models/users to select a shade/color of a lipstick and visualize how the color will suit their complexion and based on the same decide the lipstick color they wish to apply or purchase. For this application the system receives an image of the model/user and detects the lip boundary and corners. Further, the system imposes the lipstick color selected by the model/user over the detected lip boundary to enable the model/user to see how the lipstick color complements their complexion in real time. This application of the present invention eliminates the need of keeping lipstick testers or doing makeup trials.

Referring to the accompanying drawings, **FIGURE 1** displays the schematic of the system **100** for lip and lip corner detection based on machine vision based techniques.

The system **100** broadly includes the following components for lip and lip corner detection:

- a face detection unit **102** having at least one haar classifier **110** to receive and detect a face region in an image;
- a lip localization unit **116** to receive a detected face region and to locate a lip area;
- a clustering unit **126** to receive data related to the lip area from the lip localization unit **116** and obtain a lip boundary; and
- a lip corner detection unit **136** having processing means **138** to receive data related to the lip boundary and extract lip corners.

Typically, the system **100** receives an image of a user which captures sixty percent of the frontal face in the image. The captured image can be a color or a gray scale image in accordance with the present invention. This image is

represented in the form of pixels and is passed to a face detection unit **102** as the face of the individual is to be located in the image prior to detection of the lips.

The face detection unit **102** includes intermediate image representation formation means **104** to receive the pixel image and further apply predefined operations on the pixels to obtain an intermediate image representation for the image. The intermediate image representation can be computed from an image using a few operations per pixel. The conversion of an image into an intermediate image representation is described hereinafter.

In accordance with the present invention, rectangular features for facial feature extraction can be computed very rapidly using an intermediate representation of the image called the Integral Image. The Integral Image at a location x, y that is, at pixel coordinate x, y contains the sum of pixels above and to the left of x, y , inclusive:

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y')$$

where $ii(x, y)$ is the integral image and $i(x, y)$ is the original image.

Using the following pair of recurrences:

$$s(x, y) = s(x, y - 1) + i(x, y)$$

$$ii(x, y) = ii(x - 1, y) + s(x, y)$$

(where $s(x, y)$ is the cumulative row sum, $s(x, -1) = 0$, and $ii(-1, y) = 0$) the Integral Image can be computed in one pass over the original image.

Using the intermediate image representation, feature extraction means **106** of the face detection unit **102** performs feature extraction very rapidly at many scales or location in constant time.

The feature extraction means **106** receives the intermediate image representation and creates sub windows of a predefined size over the image using window creation means **108** and extracts predetermined features for the image from each of the sub windows. To detect the individual's face, Haar like features proposed by Viola and Jones [published in Second International Workshop on Statistical and Computational Theories of Vision – Modeling, Learning, Computing, And Sampling Vancouver, Canada July 13, 2001] are utilized by the feature extraction means **106**. Features are generated using a series of basic Haar Functions.

These extracted features for each of the sub windows are given as inputs to at least one haar classifier **110**. The haar classifier **110** includes the following components:

- a features repository **112** to store a set of predetermined features extracted using Haar functions;
- a training repository **114** to store a set of positive and negative images, wherein a positive image indicates a face region and negative image indicates a non-face region; and
- classification means **116** co-operating with the features repository **112** and the training repository **114** to receive the extracted features for each sub-window of the image and classify the sub-windows as positive image or negative image.

Given a features repository **112** and a training repository **114** of positive and negative images, AdaBoost is used for both to select a small set of features and to train the classifier **110**. In its original form, AdaBoost is used to boost the classification performance of a weak learning algorithm. The weak learning algorithm is designed to select a single rectangle feature which best separates the positive and negative examples. For each feature, the weak learner determines the optimal threshold classification function such that a minimum number of examples are misclassified.

The overall form of the detection process is that of a Degenerate Decision Tree, which is referred as “cascade” and is displayed in **FIGURE 2**.

All sub-windows **200** are passed to a series of cascaded classifiers **202 204 206**. A cascading haar classifier **110** achieves increased detection performance while radically reducing computation time. A key insight is that smaller boosted classifiers **110** are used which reject many of the negative sub-windows **210** while detecting almost all positive instances **208** (i.e. threshold of the boosted classifier can be adjusted so that the false negative rate is close to zero). Simple classifiers are used to reject the majority of sub-windows **210** before more complex classifiers are called upon to achieve low false positive rates.

The classified positive image or the detected facial image is passed to a Lip localization unit **118**. The Lip localization unit **118** firstly locates the central position of face through horizontal and vertical projections. The lip localization unit **118** includes histogram generation means (not shown in the figures) which receives the positive image and generates an operative vertical histogram and an operative second horizontal histogram, wherein

the vertical histogram and second horizontal histogram represent intensity values in operative horizontal and vertical directions. In accordance with the present invention, given a face image $I(i, j)$, the mean intensity for every row of the positive image is computed using equation (1) below by the histogram generation means.

$$P[i] = \sum_j M(i, j) \text{ ----- Equation (1)}$$

Further, filter application means (not shown in the figures) of the Lip localization unit **118** removes speckle noises from the vertical histogram and the second horizontal histogram and passes the filtered histogram values to computational means (not shown in the figures). The computational means receives the vertical histogram and computes an average of the vertical intensity values and marks peaks and valleys in it. The valleys (dips) in the plot of the horizontal values indicate intensity changes. After obtaining the horizontal average data, the most significant valleys (Vs) are marked. The valleys are found by finding the change in slope from negative to positive. And peaks are found by a change in slope from positive to negative.

The computational means after detecting peaks and valley obtains the center of the face as the peak value in the vertical histogram of the positive image. The peak in this graph is the possible center (PC) of the face.

The Lip localization unit **118** then obtains pixel change by performing horizontal Sobel Masking operation. The lip localization unit **118** includes binarization means **120** which converts the positive image received from the face detection unit **102** into a gray scale image. Masking means **122** of the

lip localization unit **118** receives and applies an operative horizontal Sobel mask to the gray scale image to provide a masked image. In accordance with the present invention, if I is gray scale image and G is the output image then,

$$G = I \otimes M \text{ ----- Equation (2)}$$

where M is a sobel mask / filter.

The Lip localization unit **118** further obtains a horizontal projection to locate edge of the lips. Histogram generation means **124** of the lip localization unit **118** generates an operative first horizontal histogram for the masked image, wherein the first horizontal histogram represents pixel intensity values on an operative horizontal axis. In accordance with the present invention, the lip edges are identified by computing horizontal projection on the edge mask using equation (1). The lip localization unit **118** includes lip detection means **126** which identifies the peaks in the first horizontal histogram and marks the peak as the lip area. The peaks are identified by a change in slope from positive to negative in this graph. This peak indicates the lip area.

On detecting the lip area, the data related to the lip area is passed to the clustering unit **128**. The clustering unit **128** employs lip edge dilation means **130** to perform a horizontal dilation on the detected lip area. The horizontal dilation gives a compact area where lips will be present. Thus, at this stage, the problem of lip detection is solved from a geometric perspective, where candidate lip region has been identified and includes candidate points which can be grouped into clusters for further localization.

The clustering unit **128** includes clustering means **132** which on performing a clustering operation obtains a bounding box of the lip edge giving a

definitive lip boundary. In accordance with the present invention, the input to the clustering process is expressed as the set of all candidate feature points resulting from feature segmentation. Clustering method has been implemented to select lip pixels / region. The clustering operation aims at minimizing an objective function (J), a squared error function,

$$J = \sum_{j=1}^k \sum_{i=1}^n |x_i^{(j)} - c_j|^2$$

Where $|x_i^{(j)} - c_j|^2$, a chosen distance measure between a data point $x_i^{(j)}$ and the cluster centre c_j is an indicator of the distance of the n data points from their respective cluster centers. Let the cluster centers be $c_1, c_2 \dots c_m$. They are pre-assigned as $\{c_j\}_{j=1}^m$. After calculating the objective function (J), x_i is obtained from the set of inputs. Thus, the cluster centers or centroids are updated as:

$$\text{If } J(w^{x_i}) = 1,$$

$$\text{Then } c_w^{new} = c_w^{old} + \eta (x_i - c_w^{old})$$

Where η is a small positive learning rate.

This clustering operation stops either when cluster centers stop moving (thus convergence is achieved) or when points move from one cluster to another. In the present invention, it is considered that $K=2$, one for foreground and another for background. If $O[i, j]$ is the output image for an image $G[i, j]$, then

$$O[i, j] = 1 \quad \text{If } G[i, j] > T$$

$$= 0, \text{ otherwise.}$$

T is highest value obtained from k-means clustering centre. This cluster will indicate a definitive lip boundary containing lips and thereby lip corners can be identified.

The obtained lip boundary containing the lips is then smoothened using filtering means **134** which applies a predefined filter to points in the boundary and obtains a filtered gray scale lip boundary image. In accordance with the present invention, the definitive region obtained through the clustering process is smoothened by the application of the filter displayed in **FIGURE 3**. The filter values are computed using the Hessian matrix technique by finding eigen values for the matrix.

The filtered gray scale lip boundary image is provided to a lip corner detection unit **136** having processing means **138**. The processing means **138** performs a predefined mathematical operation over the two dimensional lip region image and accurately obtains the lip corners. The mathematical operation performed the processing means is explained in detail hereinafter.

Let this image be given by I . Consider taking an image patch over the area (u,v) and shifting it by (x,y) . The weighted *sum of square difference* between these two patches, denoted S , is given by:

$$S(x, y) = \sum_u \sum_v w(u, v) (I(u, v) - I(u + x, v + y))^2$$

$$S(x, y) \approx \sum_u \sum_v w(u, v) (I_x(u, v)x + I_y(u, v)y)^2,$$

where I_x and I_y are partial derivatives of I , or equivalently

$$S(x, y) \approx \begin{pmatrix} x & y \end{pmatrix} A \begin{pmatrix} x \\ y \end{pmatrix},$$

$$A = \sum_u \sum_v w(u, v) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} = \begin{bmatrix} \langle I_x^2 \rangle & \langle I_x I_y \rangle \\ \langle I_x I_y \rangle & \langle I_y^2 \rangle \end{bmatrix}$$

-----Equation (3)

Compute $M_c = \det(A) - k * \text{trace}(A)$.

To identify best lip-corners the processing means **138** further performs Non-maximal suppression on M_c .

In accordance with the present invention, there is provided a method for detecting lips and lip corners, the method comprises the following steps as seen in **FIGURE 4**:

- detecting a face region of a user in an image, **1000**;
- localizing a lip area in the image, **1002**;
- receiving data related to the lip area and clustering the data to obtain a lip boundary, **1004**; and
- mathematically processing data related to the lip boundary and extracting lip corners, **1006**.

The method described in **FIGURE 4** further includes the steps as seen in **FIGURE 5** for lip and lip corner detection:

- capturing an image of a user, wherein the captured image is represented in the pixel form, **1000a**;
- creating sub windows of a predefined size over the image, **1000b**;
- extracting predetermined features for the image from each of the sub windows, **1000c**;

- classifying the sub windows as positive image or negative image, wherein a positive image indicates a face region and negative image indicates a non-face region, **1000d**;
- converting the positive image into a gray scale image, **1002a**;
- masking the gray scale image by applying an operative horizontal Sobel mask, **1002b**;
- generating an operative first horizontal histogram for the masked image, wherein the first horizontal histogram represents pixel intensity values on a horizontal axis, **1002c**;
- identifying peaks in the histogram and marking the peak as the lip area, **1002d**;
- performing an operative horizontal dilation on the lip area to obtain a candidate lip region, wherein the candidate lip region includes candidate points representing the candidate lip region, **1004a**;
- clustering candidate points to obtain a definitive region containing the lips, **1004b**;
- applying a predefined filter to the points in the definitive region to obtain a filtered gray scale two dimensional lip region image, **1004c**; and
- performing a predefined mathematical operation over the two dimensional lip region image to obtain the lip corners, **1006a**.

TECHNICAL ADVANTAGES

The technical advancements of the present invention include in providing a system and method for lip and lip corner detection which is based on computer vision techniques.

The proposed system rapidly detects the lip corners from an image and utilizes the determined lip localization for detecting lip corners for a makeup model company for advising lipstick colors suiting their complexion or providing makeup consultation to customers.

The proposed system captures a color / gray scale image as an input and rapidly detects the face, lips and the lip corners based on simple computational techniques making the system independent of conventional techniques like binarization, color modeling, template matching and contour extraction. The simple computer vision based techniques enable the system to accurately determine the candidate regions in real time.

The system envisaged by the present invention uses an intermediary representation of an image in the form of an integral image for processing. Hence, features used for face and lip corner detection can be very rapidly obtained at any scale or location in constant time using a few operations per pixel.

Moreover, the system employs a series of cascaded haar classifiers including at least one simple classifier and at least one complex classifier arranged in a predetermined format. The simple classifiers rapidly reject non-facial images from the intermediary representation of an image, accept and forward the candidate positive images to the subsequent classifier. The complex classifiers accept the candidate positive images from the simple classifiers, verify and reject the candidate positive images in the event that they are false positive images. Hence, these cascades of classifiers ensure that the false

positive image ratio is low and thus significantly reduce the computation time involved in processing false positive images.

The system involves minimum hardware and processing power to perform the lip and lip corner detection making the system very economical. Also, the system does not employ neural networks or supervised learning techniques for lip localization thus reducing the complexity and the training required to implement these techniques.

While considerable emphasis has been placed herein on the components and component parts of the preferred embodiments, it will be appreciated that many embodiments can be made and that many changes can be made in the preferred embodiments without departing from the principles of the invention. These and other changes in the preferred embodiment as well as other embodiments of the invention will be apparent to those skilled in the art from the disclosure herein, whereby it is to be distinctly understood that the foregoing descriptive matter is to be interpreted merely as illustrative of the invention and not as a limitation.

CLAIMS:

1. A system for lip and lip corner detection comprising:
 - a face detection unit having at least one haar classifier adapted to receive and detect a face region in an image;
 - a lip localization unit adapted to receive a detected face region and further adapted to locate a lip area;
 - a clustering unit adapted to receive data related to said lip area from said lip localization unit and obtain a lip boundary; and
 - a lip corner detection unit having processing means adapted to receive data related to said lip boundary and extract lip corners.
2. The system as claimed in claim 1, wherein said face detection unit includes a plurality of cascaded haar classifiers.
3. The system as claimed in claim 1, wherein the face detection unit includes:
 - intermediate image representation formation means adapted to receive said image in the pixel form and further adapted to apply predefined operations on pixels of the image and still further adapted to obtain an intermediate image representation for the image;
 - feature extraction means having window creation means adapted to receive said intermediate image representation and further adapted to create sub windows of predefined sizes over said image to extract predetermined 'Haar like features' from said sub windows and still further adapted to provide extracted features for each of said sub windows; and

- a plurality of cascaded haar classifiers arranged in predetermined format adapted to receive said extracted features and classify the sub windows as a positive image or a negative image, wherein a positive image indicates a face region and negative image indicates a non-face region.
4. The system as claimed in claim 3, wherein said intermediate image representation formation means is adapted to extract an intermediate representation of an image using the Integral Image technique.
5. The system as claimed in claim 1, wherein said cascaded haar classifiers are selected from the group of classifiers consisting of:
- at least one simple classifier adapted to reject sub-windows containing said negative images and further adapted to forward the candidate positive images to a succeeding classifier; and
 - at least one complex classifier adapted to receive and reject said candidate positive image in the event that the image is a false positive image.
6. The system as claimed in claim 1, wherein said cascaded haar classifiers include:
- a features repository adapted to store a set of predetermined features;
 - a training repository adapted to store a set of positive and negative images, wherein a positive image indicates a face region and negative image contains a non-face region; and

- classification means co-operating with said features repository and said training repository adapted to receive said extracted features for each sub-window of said image and further adapted to reject the sub-windows containing said negative images and accept the sub-windows containing said positive images.
7. The system as claimed in claim 1, wherein said face detection unit further includes:
- histogram generation means adapted to receive said positive image and generate an operative vertical histogram and an operative second horizontal histogram, wherein said vertical histogram and the second horizontal histogram represent intensity values in operative horizontal and vertical directions;
 - filter application means adapted to remove speckle noises from said vertical histogram and said second horizontal histogram; and
 - computational means adapted to receive said vertical histogram and compute an average of the vertical intensity values and further adapted to mark peaks and valleys in the vertical histogram and still further adapted to detect the center of the face as the peak value in the vertical histogram of said positive image.
8. The system as claimed in claim 1, wherein said lip localization unit includes:

- binarization means adapted to convert said positive image into a gray scale image;
- masking means adapted to receive and apply an operative horizontal Sobel mask to said gray scale image and further adapted to provide a masked image;
- histogram generation means adapted to generate an operative first horizontal histogram for said masked image, wherein the first horizontal histogram represents pixel intensity values on an operative horizontal axis; and
- lip detection means adapted to receive said first horizontal histogram and identify peaks in said first horizontal histogram and further adapted to mark the peaks as the lip area.

9. The system as claimed in claim 1, wherein said clustering unit includes:

- lip edge dilation means adapted to perform operative horizontal dilation on said lip area and further adapted to obtain a candidate lip region, wherein said candidate lip region includes candidate points representing the candidate lip region;
- clustering means adapted to apply a predetermined clustering technique on said candidate points and further adapted to obtain a lip boundary; and

- filtering means adapted to apply a predefined filter to points in said lip boundary and further adapted to obtain a filtered gray scale two dimensional lip region image.

10. The system as claimed in claim 1, wherein said clustering means performs the clustering operation using k-means clustering technique.

11. The system as claimed in claim 1, wherein said processing means of said lip corner detection unit is adapted to perform a mathematical operation of 'weighted sum of square difference' over said lip boundary and still further adapted to perform non-maximal suppression technique to identify lip corners.

12. A method for detecting lips and lip corners, said method comprising the following steps:

- detecting a face region of a user in an image;
- localizing a lip area in said image;
- receiving data related to said lip area and clustering the data to obtain a lip boundary; and
- mathematically processing data related to said lip boundary and extracting lip corners.

13. The method as claimed in claim 12, wherein the step of detecting a face of a user in said image includes the following steps:

- forming an intermediate image representation for said image;
- creating sub windows of a predefined size over said image;
- extracting predetermined features for said image from each of said sub windows; and

- classifying the sub windows as positive image or negative image, wherein a positive image indicates a face region and negative image indicates a non-face region.
14. The method as claimed in claim 13, wherein the step of forming an intermediate image representation for the original image includes the steps of:
- selecting pixel coordinates in the original image, one at a time; and
 - computing a value for each one of the corresponding pixel coordinates in the intermediate image representation as the sum of pixels located in the pixel coordinates above and to the left of a current pixel whose value is being computed.
15. The method as claimed in claim 13, wherein the step of classifying the sub windows as positive image or negative image includes the step of verifying the sub-windows by passing them through a series of cascading haar classifiers.
16. The method as claimed in claim 12, wherein the step of localizing a lip area in said image includes the following steps:
- converting said positive image into a gray scale image;
 - masking said gray scale image by applying an operative horizontal Sobel mask;
 - generating an operative first horizontal histogram for the masked image, wherein the first horizontal histogram represents pixel intensity values on a horizontal axis; and

- identifying peaks in said histogram and marking the peak as the lip area.

17. The method as claimed in claim 12, wherein the step of receiving data related to said lip area and clustering the data to obtain a lip boundary includes the steps of:

- performing an operative horizontal dilation on said lip area to obtain a candidate lip region, wherein said candidate lip region includes candidate points representing the candidate lip region;
- performing clustering operation on the candidate lip region to obtain a lip boundary; and
- applying a predefined filter to the points in said lip boundary to obtain a filtered gray scale lip boundary image.

18. The method as claimed in claim 12, wherein the step of mathematically processing data related to said lip boundary and extracting lip corners includes the following steps:

- creating an original image patch of a predetermined size;
- shifting the patch by a predefined factor to obtain a new shifted patch; and
- computing the weighted sum of square difference between said original image patch and said new shifted patch to identify the similarity between the two patches.

1/5

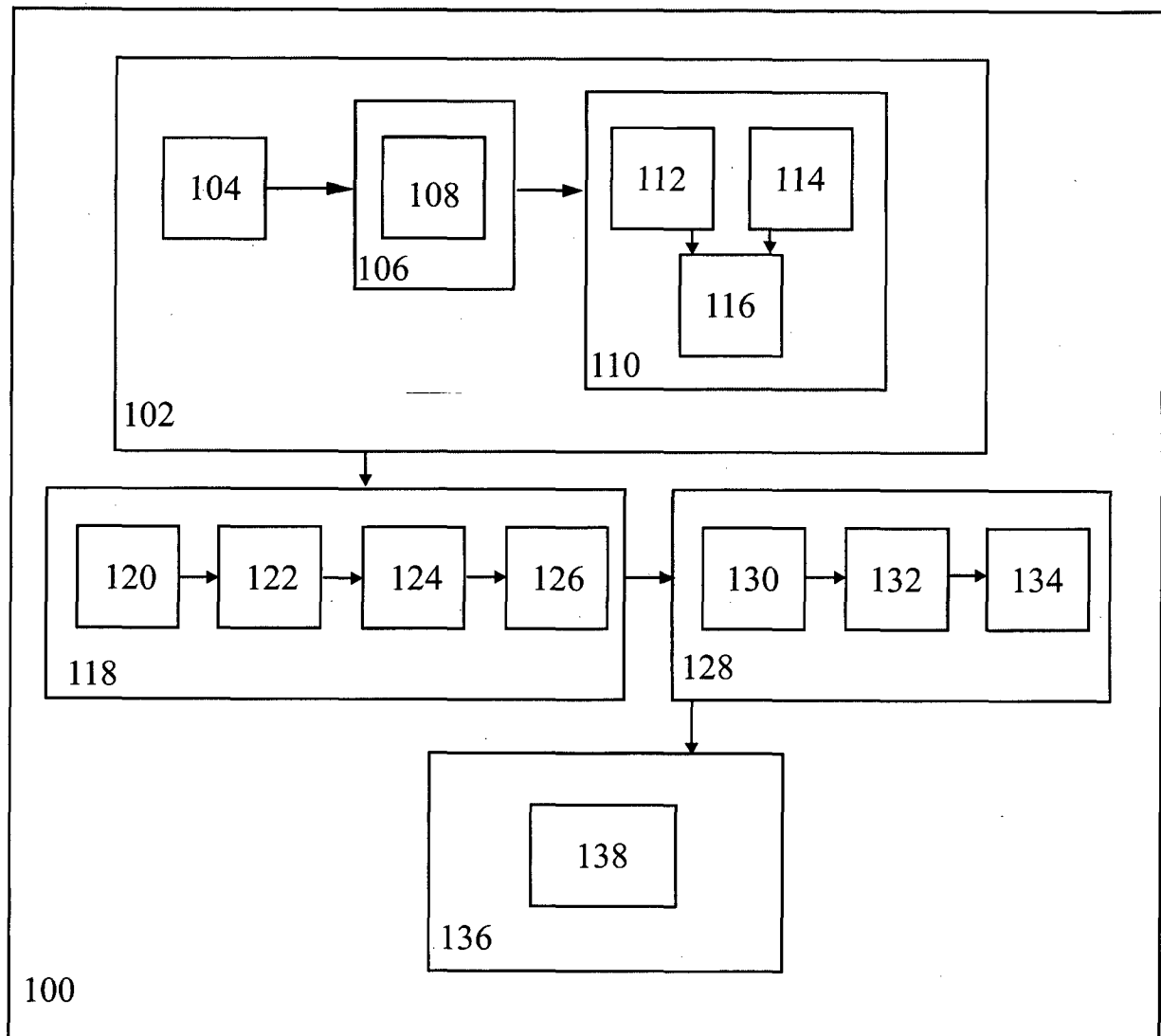


FIGURE 1

2/5

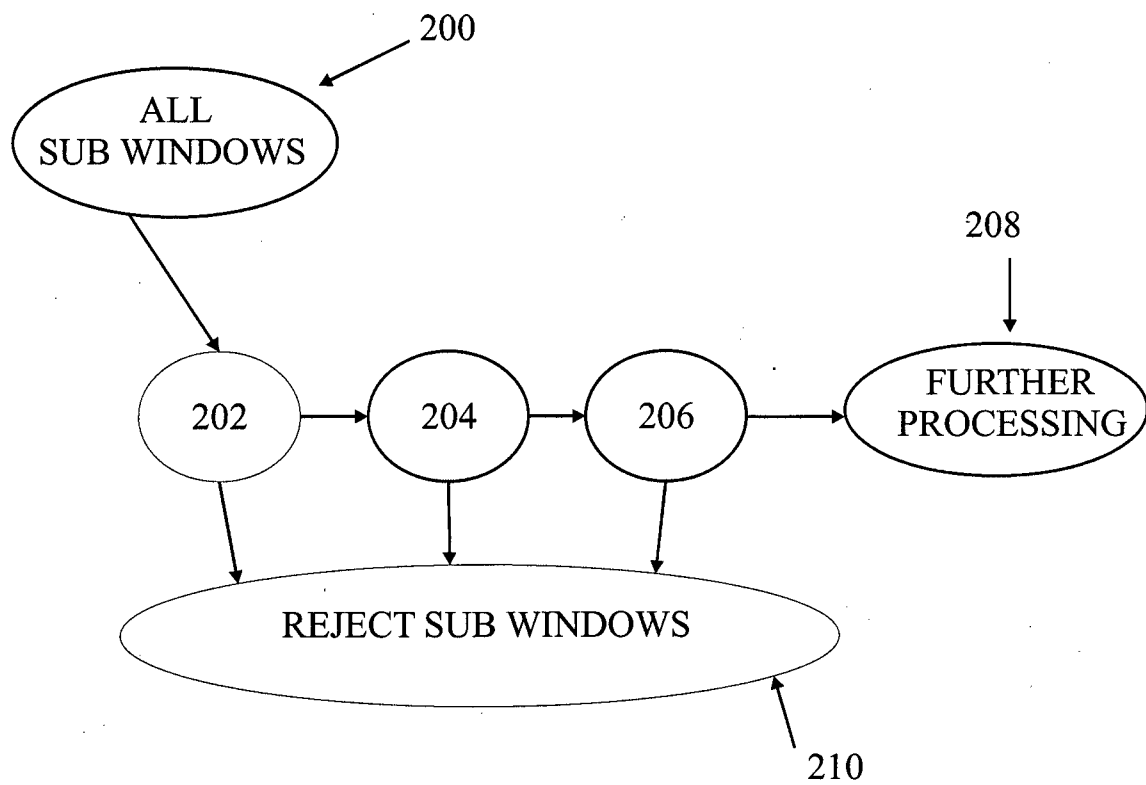


FIGURE 2

3/5

$\begin{array}{r} 1 \\ \hline 273 \end{array}$	1	4	7	4	1
	4	16	26	16	4
	7	26	41	26	7
	4	16	26	16	4
	1	4	7	4	1

FIGURE 3

4/5

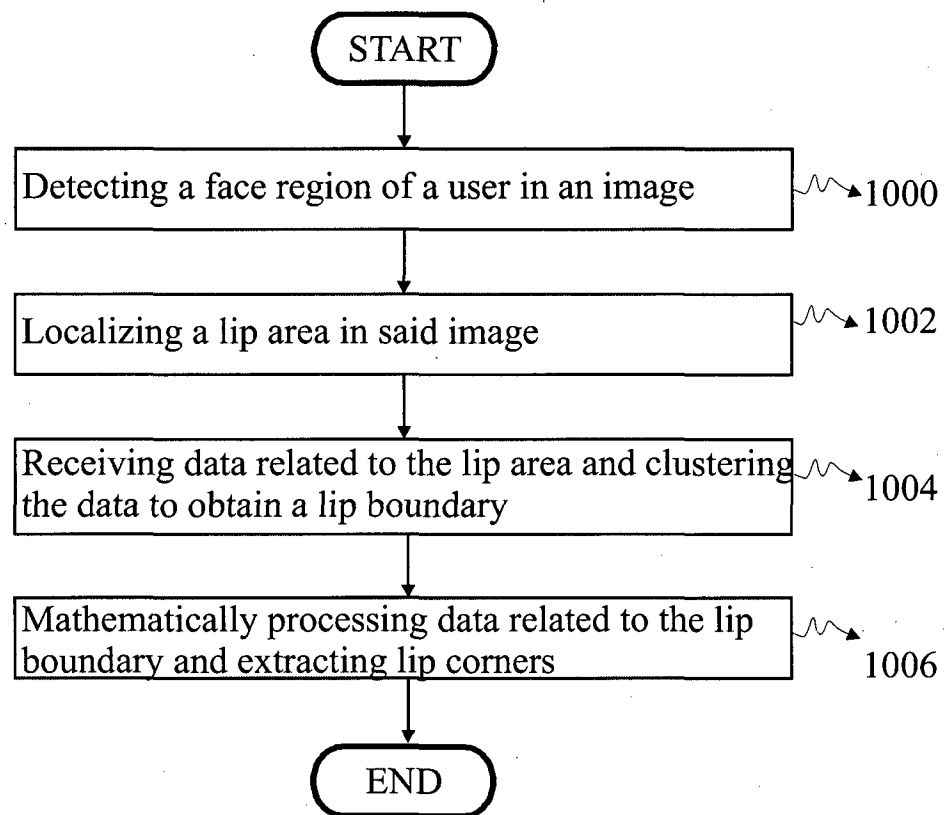


FIGURE 4

5/5

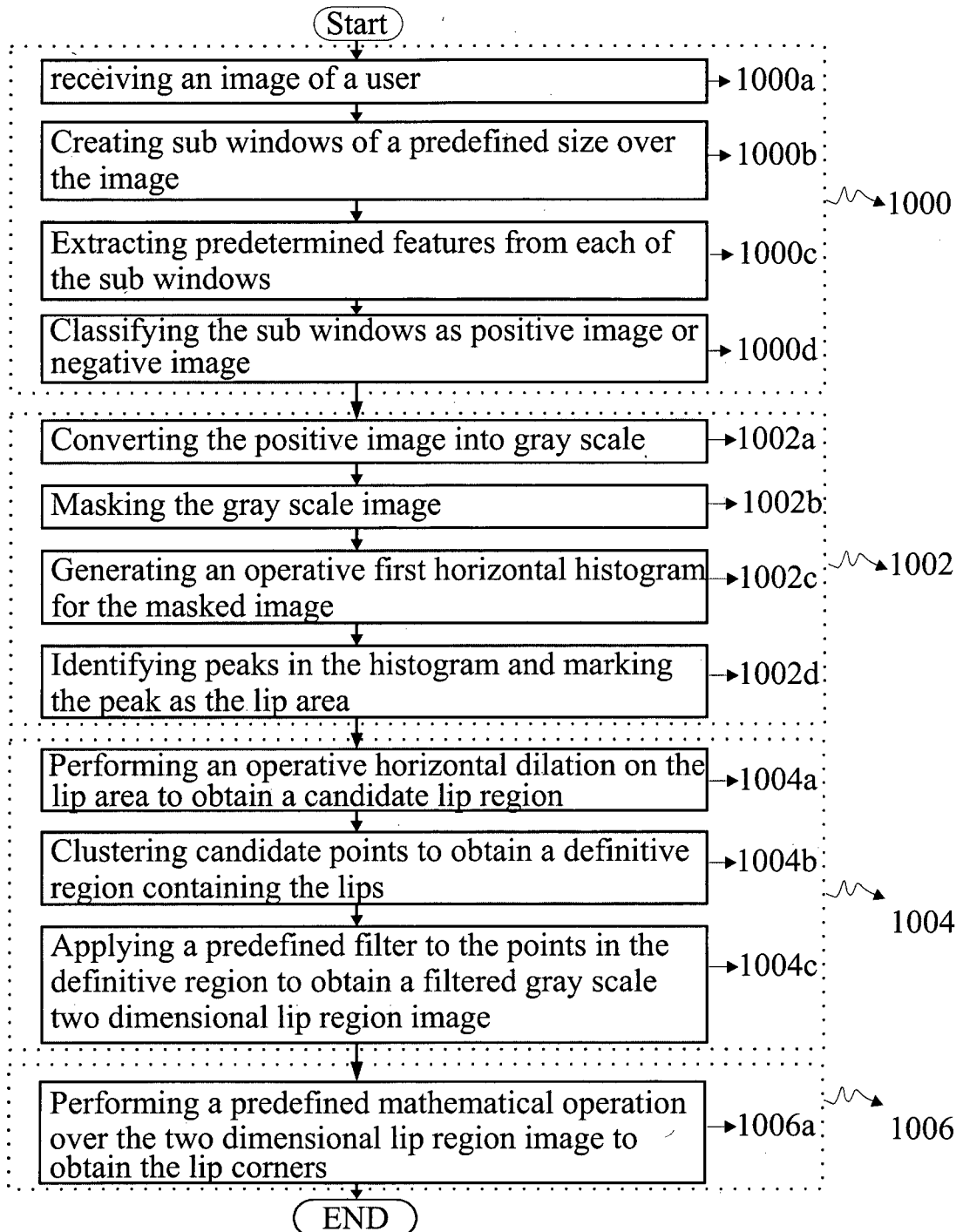


FIGURE 5