PATENT SPECIFICATION

(11) S84400

(21) Application No. S2006/0564

(22) Date of Filing of Application: 28/07/2006

(30) Priority Data:
(31) 60/746,363 (32) 03/05/2006 (33) United States of America (US)

(45) Specification Published: 1 November 2006

(51) Int. Cl. (2006)
G06T 5/50
G06T 7/00

---

(54) Title: Improved foreground / background separation

(72) Inventor: MIHAI CIUC

(73) Patent Granted to: FOTONATION VISION LIMITED, An Irish company, Galway Business Park, Dangan, Galway, Ireland

Improved Foreground / Background Separation

The present invention provides a method and apparatus for providing improved foreground / background separation in a digital image.

It is known to build a focus map using a depth from defocus (DFD) algorithm, for example, as disclosed in "*Rational Filters for Passive Depth from Defocus*" by Masahiro Watanabe and Shree K. Nayar (1995). The basic idea is that a depth map of a given scene can be theoretically computed from two images of the same scene. Ideally, for calculating a DFD map, a telecentric lens is used, and only focus varies between the two image acquisitions. This is generally not true of existing digital cameras.

Another technique for separating foreground from background is disclosed in as described in our co-pending Irish Patent Application No. S2005/0822 filed December 8, 2005. Here, the difference in exposure levels between flash and non-flash images of a scene are used to provide a foreground/background map. The main advantage of using depth from defocus over a flash/non-flash based technique, is that it depth from defocus is independent of the scene illumination and so can be more useful for outdoor or well-illuminated scenes.

A further technique for separating foreground from background is disclosed disclosed in US Application No. 60/773,714. Here, the difference in high frequency coefficients between corresponding regions of images of a scene taken at different focal lengths are used to provide a foreground/background map. Again in this case, the foreground/background map is independent of the scene illumination and so this technique can be useful for outdoor or well-illuminated scenes.

In any case, the foreground/background map produced by each of the above techniques or indeed any other technique may not work correctly. It is thus desirable to provide an improved method of foreground/background separation in a digital image.

According to the present invention there is provided a method as claimed in claim 1.

Embodiments of the invention will now be described with reference to the accompanying drawings, in which:

Figure 1(a) shows an in-focus image of a subject; Figure 1(b) shows a DFD map for the image; and Figure 1(c) shows the DFD map of Figure 1(b) partially processed according to a preferred embodiment of the invention;

Figure 2 shows a flow diagram of a method for improving foreground/background separation according to the preferred embodiment of the invention;

Figure 3(a) shows a first colour segmented version of the foreground regions of the image of Figure 1(c); Figure 3(b) shows a profile for a subject; Figure 3(c) shows the result of combining the profile of Figure 3(b) with the regions of Figure 3(a) according to an embodiment of the present invention; and Figure 3(d) shows the image information for the identified foreground regions of the image of Figure 3(c);

Figure 4(a) shows another in-focus image of a subject; Figure 4(b) shows a DFD map of the image; Figure 4(c) shows a first color segmented version of the foreground regions of the image; and Figure 4(d) shows the result of combining a profile with the regions of Figure 4(c) according to an embodiment of the present invention;

Figure 5(a) shows another in-focus image of a subject; Figure 5(b) shows a first color segmented version of the foreground regions of the image; and Figure 5(c) shows a further improved color segmented version of the foreground regions of the image when processed according to an embodiment of the present invention; and

Figure 6(a) – (c) show luminance histograms for regions identified in Figure 5(a).

The present invention is employed where there is a need for foreground/background segmentation of a digital image. There are many reasons for needing to do so, but in particular, this is useful where one of the foreground or the background of an image needs to be post-processed separately from the other of the foreground or background. For example, for red-eye detection and correction, it can be computationally more efficient to only search and/or correct red-eye defects in foreground regions rather than across a complete image. Alternatively, it may be desirable to apply blur only to background regions of an image. Thus, the more effectively foreground can be separated from background, the better the results of image post-processing.

In the preferred embodiment, improved foreground/background segmentation is implemented within digital camera image processing software, hardware or firmware. The segmentation can be performed at image acquisition time; in a background process, which runs during camera idle time; or in response to user interaction with image post-processing software. It will nonetheless be seen that the invention could equally be implemented off-line within image processing software running on a general-purpose computer.

In any case, in the preferred embodiment, a user operating a camera selects, for example, a portrait mode and optionally a particular type of portrait mode, for example, close-up, mid-shot, full length or group. In portrait mode, the camera then acquires a main image or indeed the camera acquires one of a sequence of preview or post-view images generally of the main image scene. Generally speaking, these preview and post-view images are of a lower resolution than the main image. As outlined above, at some time after image acquisition, image processing software calculates either for the main image or one of the preview/post-view images an initial foreground/background map.

The preferred embodiment will be described in terms of the initial map being a DFD map, although it will be appreciated that the invention is applicable to any form of initial foreground/background map as outlined above. In the embodiment, the segmentation process provides from the initial map, a final foreground/background map, where the foreground region(s), ideally, contain the image subject and which can be used in further image processing as required.

Figure 1(a) shows an in-focus image of a scene including a subject (person) 10 and Figure 1 (b) the resulting DFD map. The DFD map has, in general, a number of problems in that:

- objects such as the shutters 12 that lie in the neighborhood of the subject although at different depths appear in-focus (which is normal, but undesired) and as such can be falsely classified as foreground objects; and

- the DFD map is very noisy, i.e., it is far from being smooth.

Referring now to Figure 2, the foreground/background segmentation processing of the DFD map to provide the final foreground/background map is shown:

The initial DFD map 20, for example, as shown in Figure 1(b), is first smoothed or blurred with a Gaussian kernel, step 22. The DFD map of Figure 1(b) is in a binary form with white regions being classified as foreground and black being background. Smoothing/blurring the map will tend to indicate foreground regions as generally lighter and background regions as generally darker.

A threshold is then applied, step 24, to the smoothed continuously valued image from step 22. This provides a binary map in general having larger and smoother contiguous regions than the initial DFD map 20.

Regions of the binary map obtained at step 24 are then filled, step 26, to remove small regions within larger regions. For the initial image of Figure 1(a), an initial foreground/background map as shown in Figure 1(c) is produced. Here foreground is shown as white and background as black. It will be seen that in this map, there is no distinction between the foreground subject region 14 and the region 16 which should be in the background.

The pixels classified as background in the image of Figure 1(c) are excluded from further processing, step 28, and the remaining regions of the images are regarded as provisional foreground regions.

The remainder of the image is segmented by color, using any suitable technique, step 30. In the preferred embodiment, a *"mean shift" algorithm*, based on D. Comaniciu & P. Meer, "Mean Shift: A Robust Approach toward Feature Space Analysis" IEEE Trans. Pattern Analysis Machine Intell., Vol. 24, No. 5, 603-619, 2002) is employed. In general, this technique involves identifying discrete peaks in colour space and segmenting the image into regions labelled according to their proximity to these peaks.

While this technique can be performed in RGB space, for the sake of computational complexity, the preferred embodiment operates on [a,b] parameters from an LAB space version of the foreground region 14,16 pixels. This means that for an image captured in RGB space, only pixels for candidate foreground regions need to be transformed into LAB space. In any case, it should be noted that this [a,b] based segmentation is luminance (L in LAB space) independent. This segmentation produces a map as shown in Figure 3(a), where the different shaded regions 30(a)..30(f) etc represent a region generally of a given [a,b] colour combination.

In a first improvement of foreground/background segmentation according to the present invention, a portrait template corresponding to the acquired image is provided, Figure 3(b). The template includes a profile 32 of a subject. The exact size of a particular profile can be varied according to the focal length for the acquired image in accordance with the expected size of a subject. It will be seen that while the profile 32 shown in Figure 3(b) is a mid-shot of a subject, the outline can be varied according to the expected pose of a subject. This can either entered manually by a user, by selecting a suitable portrait mode, or possibly predicted by the image processing software. Thus, the profile might be a head shot outline or a full body outline, in one of a plurality of poses, or indeed in the case of a group portrait, an outline of a group.

In any case, the color segments provided in step 30 are combined with the profile 32 to retain only color regions that overlap to a significant extent with the profile 32. Thus, with reference to Figure 3(a), it will be seen that inter alia regions 30(b),(c) and (e) are removed from the foreground map, while inter alia regions 30(a),(d) and (f) are retained. The final set of foreground regions is shown shaded in Figure 3(c), with the final background region being indicated as black. It will been seen, however, from Figure 3(d)) that some regions such as sub-region 30(g) of region 30(a) are still not as accurately segmented as they might be.

It will be seen that sub-regions 30(g)(1) and 30(g)(2), because they may have similar [a,b], characteristics have been included in region 30(a) which in turn has been classified as a foreground region, whereas sub-region 30(g)(2) should more suitably be classed as a background.

It is also acknowledged that parts of the foreground can be (wrongly) removed from the foreground map from various reasons. For instance, in Figure 3(d), it can be seen that the subject's right hand has been removed from the foreground map because it does not overlap with portrait profile 32.

Another example of the segmentation of steps 22-34 is illustrated with reference to Figure 4. Figure 4(a) shows an in-focus image and Figure 4(b) the DFD map for the image. Figure 4(c) shows the segmented map after color segmentation, step 30. Figure 4(d) shows the final foreground/background map after elimination of regions, such as 40(a),(b) that do not overlap significantly to a portrait template 32 chosen for the image.

It can be seen in this case that, because color segmentation did not separate the subject's hair from the balcony's edges, region 40(c), the balcony edges have been wrongly included in the final map as foreground regions.

In a still further example, Figure 5(a) shows an in-focus image of a subject and Figure 5(b), the foreground/background map after color segmentation, step 30, but before combining the foreground regions with a profile 32. Two segmentation artifacts can be seen at this stage: the subject's T-shirt 50 and the TV 52 behind are segmented in a single region; and, similarly, half the subject's face and hair 54 are merged into a single region. The latter defect (accidentally) will not affect the final results, as both hair and face are ideally included in a final foreground map. On the contrary, not separating the T-shirt 50 from the TV 52 results in (wrongly) retaining the latter in the foreground map.

In a second improvement of foreground/background segmentation according to the present invention, foreground regions are analysed according to luminance, step 36. This step can be performed in addition to, independently of, or before or after step 34. In the preferred embodiment, this analysis is again performed on an LAB space version of the foreground region 14,16 pixels and so can beneficially use only the L values for pixels as is described in more detail below.

In step 36, the intensity of the pixels in regions of the image of interest is analysed to determine if the luminance distribution of a region is unimodal or bimodal. This, in turn, allows difficult images to have their foreground/background regions better separated by applying unimodal or bimodal thresholding to different luminance sub-regions within regions of the image.

In the case of Figure 5, both the T-shirt/TV 50/52 and hair/face pairs 54 strongly differ in luminance. In step 36, the luminance histogram of each segmented foreground region is computed. Figure 6 shows the luminance histograms of region #1 comprising the T-shirt/TV 50/52; region #2 comprising the hair/face 54; and region #3 shows a typical unimodal distribution. As can be seen from Fig. 6, the luminance histograms of regions that should be further segmented (i.e., regions # 1 and 2) are bi-modal, whereas others (region #3) are not.

It should also be noted that multi-modal histograms could also be found for a region, indicating that the region should be split into more than two regions. However, the instances of such a distribution are likely to be very rare.

Given that regions which exhibit such a bi-modal distribution in luminance should be ideally segmented further, it is useful to conveniently classify a given histogram as either unimodal or bimodal. Referring to Figure 6, in the preferred embodiment, this classification comprises:

(i) blurring/smoothing the histogram to reduce artifacts;

(ii) finding a maximum luminance 60 in the histogram;

(iii) discarding a given-width interval 62, Figure 6(a), around the maximum coordinate (to avoid detection of false maxima);

(iv) finding the next maximum 64;

(v) from each of the two maxima, a mode-detection procedure is run to find the corresponding mode – a Bell shaped distribution around each maximum, 66, Figure 6(b);

(vi-a) if both found modes include a significant portion of the histogram (i.e., if each spans an interval of luminance levels that includes more than 20% of the pixels from regions of interest) then the histogram is declared bimodal, and the minimum value 68 in the interval between the two maxima is used as a threshold for splitting the region into 2 sub-regions; otherwise,

(vi-b) the histogram is said to be unimodal, and nothing is changed.

Figure 5(c) presents the result of the final segmentation, where one can see the correct separation of T-shirt/TV and of hair/face pairs. Regions which are considered unimodal are not changed.

Using the present invention, more of an in-focus subject can be correctly separated from the background, even in difficult images, i.e., images with background located very close to the subject. Even when portions of background cannot be separated from the foreground or vice versa, the artifacts are less likely to be big, and the final map can be more useful for further post-processing of the image.

There are a number of practical issues, which need to be considered when implementing the invention:

When the initial map is derived from a DFD map, then the scaling factor between the in-focus and out-of-focus images will need to be known. This needs to be accessible from the camera configuration at image acquisition, as it cannot be computed automatically. It is derivable from knowing the focal length for the acquired image, and so this should be made available by the camera producer with the acquired image.

It will also be seen that where the initial map is derived from a DFD map, some shifting between images may have taken place, depending upon the time between acquiring the two images. It will be seen that the subject may move significantly with respect to the background, or the whole scene may be shifted owing to camera displacement. As such appropriate alignment between images prior to producing the DFD map should be performed.

As indicated earlier, the invention can be implemented using either full resolution images or sub-sampled versions of such images, such as pre-view or post-view images. The latter may in fact be necessary where a camera producer decides double full resolution image acquisition to provide a full resolution DFD map is not feasible. Nonetheless, using a pair comprising a full-resolution and a preview/postview, or even a pair of previews/postviews for foreground/background mapping may be sufficient and also preferable from a computational efficiency point of view.

It will also be seen that it may not be appropriate to mix flash and non-flash images of a scene for calculating the DFD map. As such, where the main image is acquired with a flash, non-flash preview and post-view images may be best used to provide the foreground/background map in spite of the difference in resolution vis-à-vis the main image.

Claims:

1. A method for providing improved foreground / background separation in a digital image of a scene comprising the steps of:

providing a first map comprising one or more regions provisionally defined as one of foreground or background within said digital image;

providing a subject profile corresponding to a region of interest of said digital image;

comparing at least some of said one or more provisionally defined regions with said subject profile to determine if any of said regions intersect with said profile region;

changing in said map the definition of one or more of said regions based on said comparison.

2. A method as claimed in claim 1 wherein said changing step comprises:

comparing a foreground region with said subject profile; and

responsive to said foreground region not substantially intersecting said subject profile, changing the definition of said foreground region to a background region.

3. A method as claimed in claim 1 wherein the step of providing said first map is based on a comparison of two or more images nominally of said scene.

4. A method as claimed in claim 3 wherein one or more of said two or more images is a low resolution version of said digital image.

5. A method as claimed in claim 3 wherein one of said two or more images is said digital image.

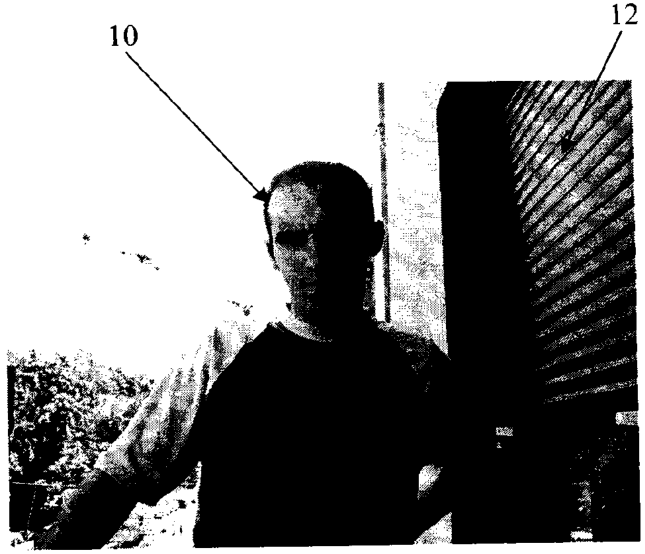


Figure 1(a)

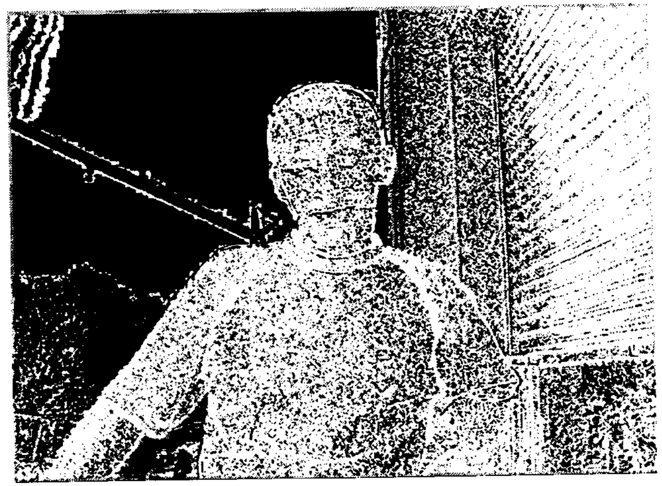


Figure 1(b)

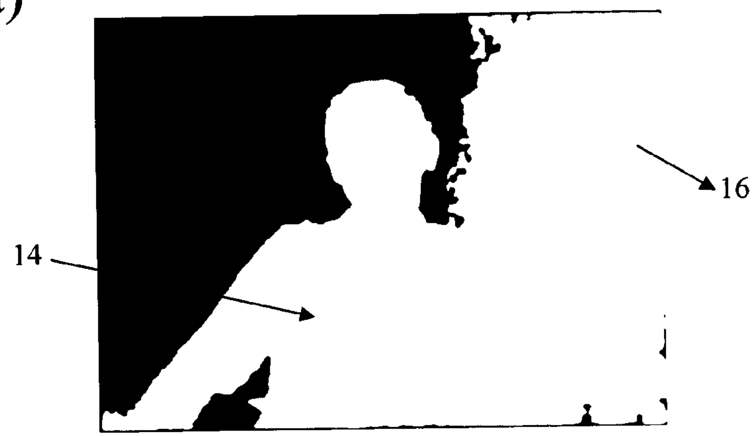
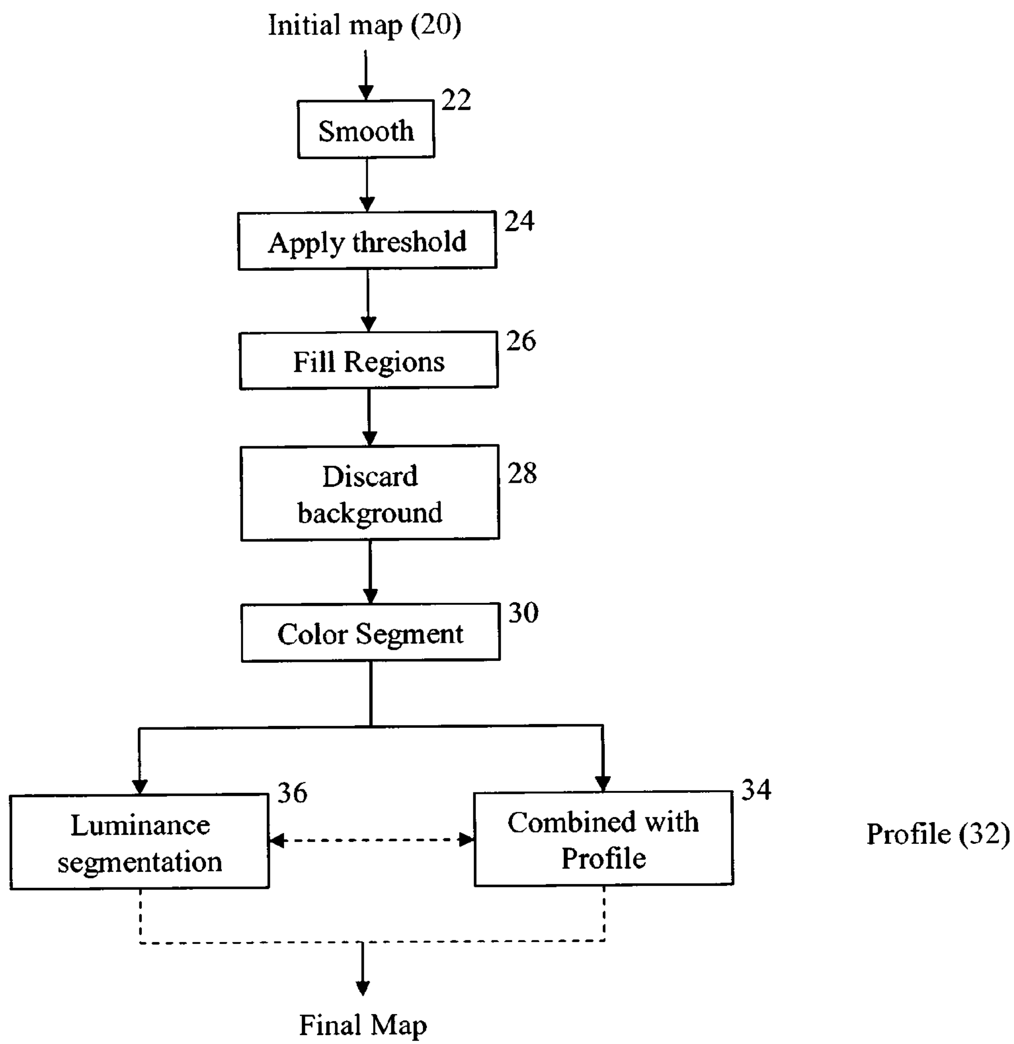


Figure 1(c)

S84400

Figure 2



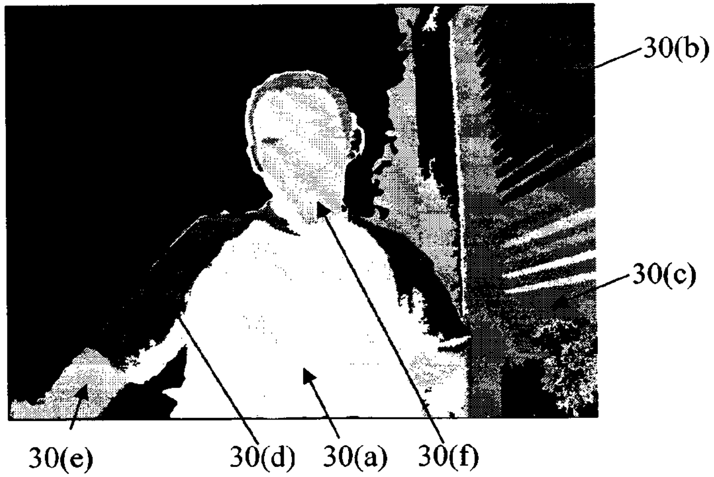


Figure 3(a)

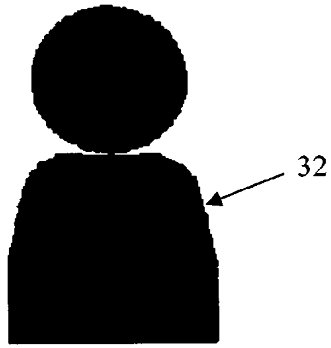


Figure 3(b)

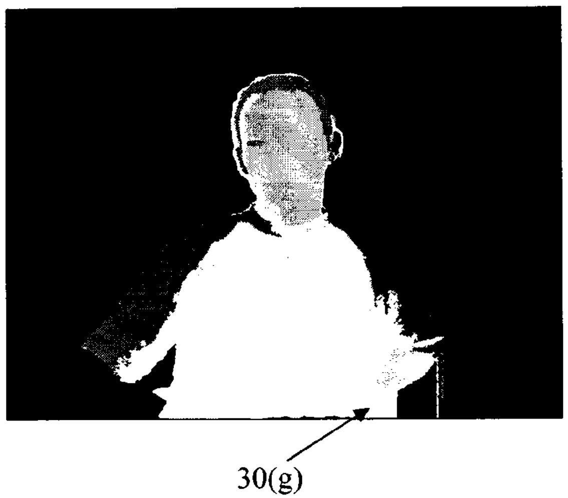


Figure 3(c)

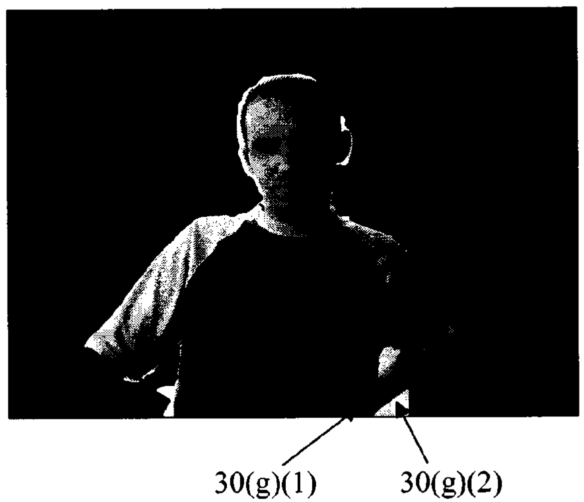


Figure 3(d)

Figure 4(a)

Figure 4(b)

Figure 4(c)

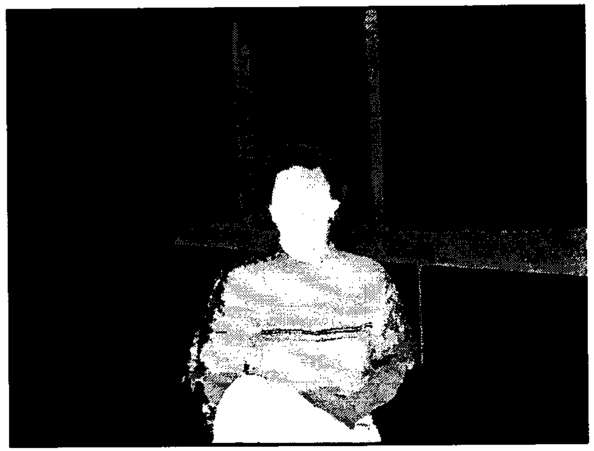


Figure 4(d)

S84400

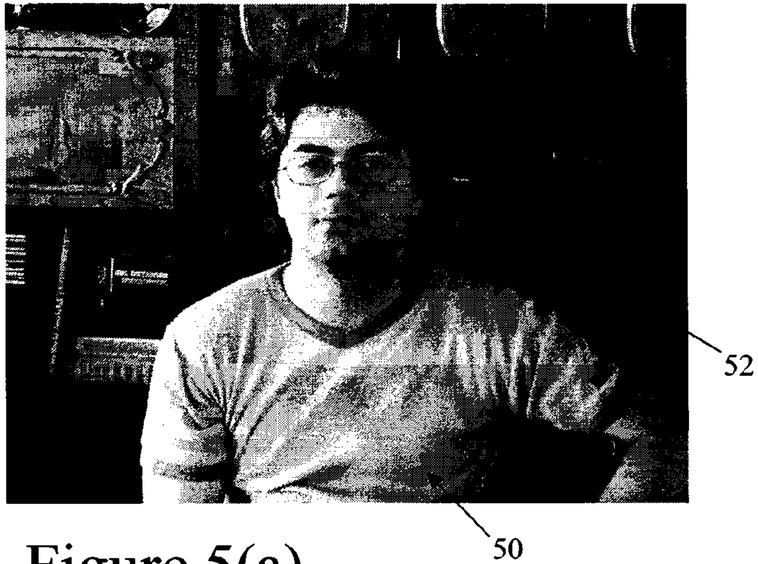


Figure 5(a)

Figure 5(b)

Figure 5(c)

S84400

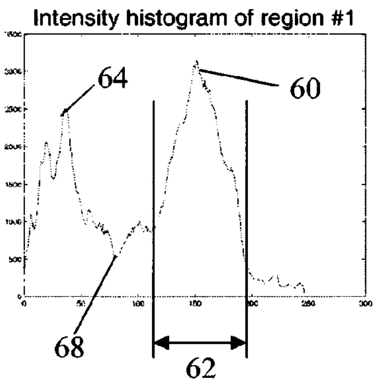


Figure 6(a)

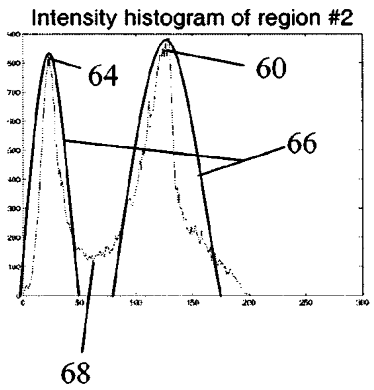


Figure 6(b)

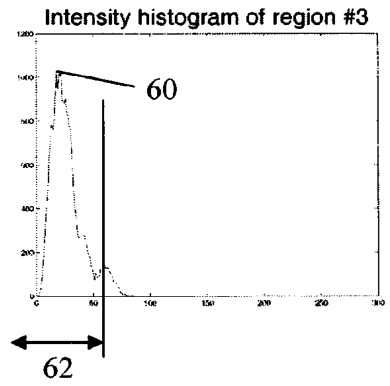


Figure 6(c)