Application of Colour Detection and Snakes To Track Hand Images

S. ROBLA, C. TORRE, E.G. SARABIA, J.R. LLATA
Electronic Technology and Automatic System Department.
University of Cantabria
Avda. Los Castros s/n, 39005 Santander
SPAIN

Abstract: - This article presents a study about two popular segmentation and tracking techniques in computer vision, skin segmentation and snakes, and their application to a particular case. The study is centered in skin segmentation over a wooden background mostly. At first wood and skin characteristics were thought to be very similar, what would have been an inconvenient. However, this paper shows that it is possible to get a satisfactory colour segmentation applying a Bayes classifier in YCbCr images format. The use of automatically initialized open active contours let the tracking of the segmented hand and its localization in the image.

Key-Words: - **:** Snakes, Tracking, Colour Segmentation, Bayes Classifier, Safety System.

1 Introduction

This work presents a combination of two techniques of colour segmentation and opens active contours to track hands.

As an application to both techniques it was developed a security system for circular saw benches. Thirty five percent of the accidents with circular saw benches result in the amputation of fingers due to the risk of being trapped by the saw while the wood is being handled. So the designed system attempts to solve this safety question. Currently existing safety methods are mechanical type. The developed system is based on artificial vision and it is a complement to the safety elements located at the top of the machine, such as a riving knife, or a saw guard among others.

One of the most interesting aspects in this application is the fact that hands change in pose (location and orientation) and shape (open and close) during the cutting tasks. A visual tracking of deformable objects algorithm has been used, in order to detect and follow the hands movements on the bench.

The artificial vision camera of the developed system is located over the working area and it captures the image of both, the saw and part of the moving table nearer the saw. These images are processed to detect the presence of the hand. Subsequently, the safety system will be activated, and in an automatic way, an open active contour will be generated for the detected hand. At this moment, the snake allows to follow its movements independently of the wood movement. This active contour is fitted with a safety margin that increases the snake around the hand to reduce the risk of contact with the saw blade. Therefore before the hand gets too close to the danger zone the system generates a warning, and the cut mechanism is stopped. If there is no hand in the images, the safety system will remain dormant, being activated by the presence of a hand.

1.1 Colour spaces used in skin segmentation

Detection of skin colour in images is an extensive topic in computer vision. In this work it will be used the detection of skin colour images, particularly for hands classification. One important factor that can influence the obtained segmentation is the variation of lightning conditions since intensity of pixels can vary.

To diminish this effect in skin segmentation it have be taken into account the actual colour spaces, since they have different behaviors have tried several colour spaces for various hand motion images sequences taken over different wood backgrounds and under slight light variations.

1.1.1 RGB colour space

The RGB is the most common format of colour images and three numerical components are used to represent a colour. Then this colour space can be represented as a three dimensional coordinate system, where the axes are the components R(Red),

G(Green), and B(Blue). Sixteen million possible colours can be generated through the combination of these primary colours. But this format is very sensible to the lighting conditions because these values about tone, saturation, and hue are mingled. Sometimes color segmentation algorithms are not robust enough due to the variation of luminance effects from the lighting conditions.
Some researches have defined a rule to define skin region [3],[5] as a possible solution to this inconvenient,

(R, G, B) is classified as skin if:

R>95 and G>40 and B>20 and
max{R,G,B}-min(R,G,B)>15 and (1)
|R-G|>15 and R>G and R>B

Another representation of colour skin is achieved using the normalized RGB space.

$$r = \frac{R}{R+G+B} \quad g = \frac{G}{R+G+B} \quad b = \frac{B}{R+G+B} \qquad (2)$$

Where r+g+b=1. The b component can be omitted since does not add any significant information. The r, g components are called “pure colours” for the dependence of r and g on the brightness of the RGB colour is decreased by the normalization.

1.1.2 HSI colour space

This colour space has the property of separating the luminance component form chromatic component and with that at least partial independence of chromaticity and luminance is achieved.
It gives the possibility of searching skin colour elements in 2D chromatic space, using only two dimensions to describe the colour.
Hue describes the basic pure colour of the image, Saturation measures the colorfulness of an area in proportion to its brightness. Indicating to what extend a pure colour is diluted with white colour. Intensity is related to the colour luminance.
Hue is invariant to highlights at white light sources, and also for matte surfaces, to ambient light and surface orientation relative to the light source.
Values of Hue and Saturation give useful information about skin characteristics.
The HIS is obtained by a no-linear transformation. Equations 3 show this colour space conversion from RGB to HSI [2]

$$H = ar\cos\frac{\frac{1}{2}((R-G)+(R-B))}{\sqrt{(R-G)^2+(R-B)(G-B)}}$$
$$S = 1 - 3\frac{\min(R,G,B)}{R+G+B} \qquad (3)$$
$$I = \frac{1}{3}(R+G+B)$$

To make the skin detection several test have been carried out to delimit the range of Hue, and Saturation by establishing a threshold. The condition expressed in 4 has defined with it has obtained better results as can bee seen in the segmentation in Fig.1and the skin distribution pixels in Fig.2

H<=0.14 & S>=0.55 (4)

Fig. 1

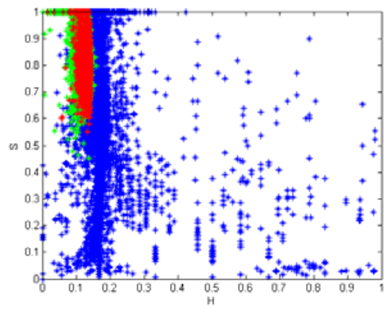

Fig. 2

However if illumination change the condition may fail.

1.1.3 YCbCr colour space

The component Y contains the luminance information, whereas the components Cb y Cr, contain the chrominance information.

$$Y = 0.299R + 0.587G + 0.114B$$
$$C_r = R - Y \qquad (5)$$
$$C_b = B - Y$$

This transformation brings the separation of luminance and chrominance components.

The skin colour segmentation algorithms become more robust using only the chrominance and the removing luminance.

1.2 Behavior of the snakes

Active contours are based in a deformable curve composed of elastic elements that provide the characteristics of rigidity, and elasticity. These elements are defined by a set of control points that move under the influence of internal or external forces.
A traditional deformable curve can be defined in the (x,y) plane [1] as a parametric curve

$$v(s,t) = (x(s,t), y(s,t)) \forall s \Omega t \in T \quad (6)$$

where s is the spatial index (x,y), t the time, Ω the spatial and T the time region, respectively.
This curve can be expressed based on energies as

$$E_{Snake}(v) = \frac{1}{2} \int_{\Omega} [E_{Int}(v) + E_{Ext}(v)] \, ds \quad (7)$$

$E_{Internal}$ keeps the curve together and prevents it from bending too much, then the snake tends to be more continuous and smooth. $E_{External}$ includes external constraints forces that located the snake near the desired local minimum. And E_{Field} gravitational field force push the snake towards salient image features.

$$E_{Internal} = E_{Elastic} + E_{Bending} \quad (8)$$

where,

$$E_{Elastic}(v) = \alpha(s) \left[\frac{dv(s)}{ds} \right]^2 \quad (9)$$

where α is the parameter that regulates the tension of the snake and

$$E_{Bending}(v) = \beta(s) \left[\frac{dv^2(s)}{ds} \right]^2 \quad (10)$$

where β is the parameter that regulates the rigidity.
The field energy [7] function is defined as:

$$E_{Ext}(v) = k \left[\nabla \left[G_{\sigma}(x,y) * I(x,y) \right] \right]^2 \quad (11)$$

where k is a weighting parameter, ∇ is the gradient operator, $G_{\sigma}(x,y)$ is a two-dimensional Gaussian function, with a standard deviation σ, and * is the 2D image convolution operator.

The snakes try to match a deformable image minimizing E_{snake}. The snake that minimize that energy [6] must satisfy the Euler equation

$$E_{Elastic} - E_{Bending} - E_{Ext} = 0 \quad (12)$$

Finding a curve x(s,t) that minimizes the energy of the snake is a variational problem, and the numerical solution arises by discretizing the equations of motion, and solving the discrete system iteratively. From this discretization a useful matrix K [1], consisting of tension, rigidity and field parameters is generated.

2 Hand colour segmentation

About 50 images have been taken to create a database that lets analyze skin regions and non-skin regions taken principally in different wood backgrounds. The distribution of the pixels of training images, for skin class (green) and for non-skin class (blue) are depicted in Fig.3 that shows the R-G, B-R, and B-G, Fig.4 that shows the CrCb pixels distribution and Fig.5 that shows the HS pixel distribution.

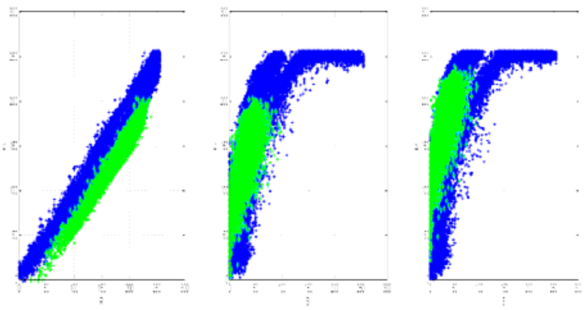

Fig 3

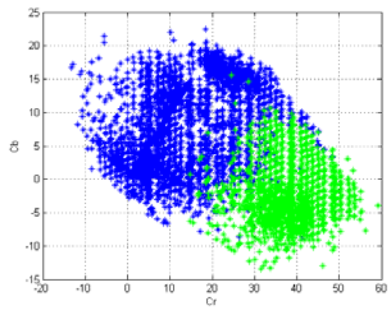

Fig.4

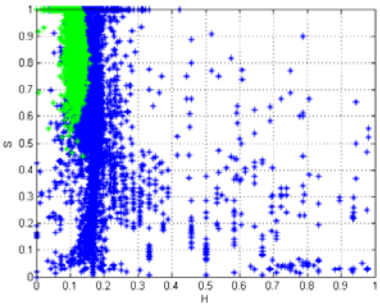

Fig.5

As it can be observed the distribution of the skin pixels and non-skin pixels is more differentiated in YCbCr format and HSV than RGB. A density probability function (pdf) in HSV, (the obtained results were not satisfactory), was defined as well as another pdf to the YCbCr format to skin region as a Gaussian model characterized by the covariance and the mean of chrominance values. Then the probability of a pixel to lies in the skin region is given by [5] :

$$\mathrm{p}(\alpha \mid \mathrm{skin}) = \frac{1}{2\pi\sqrt{\sum_s}} e^{-\frac{1}{2}(c-\mu_s)^T \sum_s^{-1} (c-\mu_s)} \qquad (13)$$

Where c is the color vector, μ_s is the mean vector given by the equation 14, and $\sum_s$ the covariance matrix given by the equation 15

$$\mu_s = \frac{1}{n}\sum_{j=1}^{n} c_j \qquad (14)$$

$$\sum_s = \frac{1}{n-1}\sum_{j=1}^{n}(c_j - \mu_s)(c_j - \mu_s)^T \qquad (15)$$

p(α|skin) represents the probability of the occurrence of a given pixel that is a skin pixel, and μ_s represent the mean vector, and Σ_s the covariance matrix of the training pixels. In a analog way p(α|nonskin) represents the probability of the occurrence of a given pixel occurring that it is in the non-skin region.

Fig.6 shows this probability density function where skin class has a minor deviation than non-skin class, and both are overlapped in a small area.

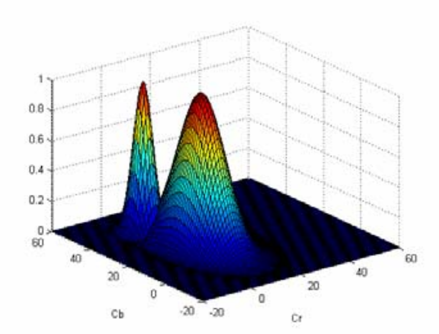

Fig.6

Fig.7 shows the probability density function of skin class and the probability density function of wood class.

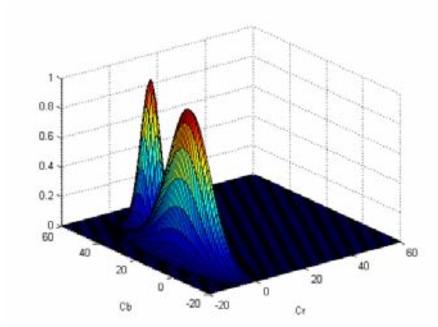

Fig.7

It can be observed that the probability of wood is within the range of the probability density of non-skin class, thus the classification is made using no skin class.

Taking into account these results have been implemented a Bayes classifier to differentiate both classes. Part of the images have been used for the training of the classifier, and the others as probe images.

$$p(skin \mid \alpha) = \frac{p(\alpha \mid skin)\, p(skin)}{p(\alpha \mid skin)\, p(skin) + p(\alpha \mid noskin)\, p(noskin)} \qquad (16)$$

p(skin| α) gives the probability that a pixel represents skin once given a chrominance value.

p(skin) is the probability of skin and non-skin pixels occurrence. In this work both have been considered to have an equal probability, so the p(skin) and the p(nonskin) have been taken as 0.5.

The obtained segmentations can be seen in Fig.8 and Fig.9 along with its skin pixels distribution in Fig.10.

Fig.8

Fig.9

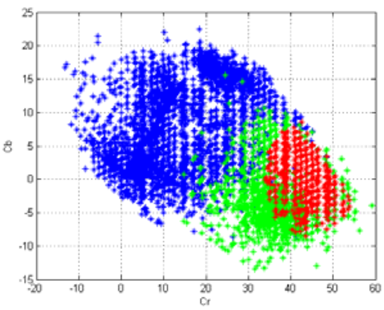

Fig.10

3 Generating snakes

Once the hand segmentation has been achieved successfully, the image is transformed into binary.
A process of filling holes and "filter" is followed to obtain the edge of the image. In this edge two processes operations are performed, one involving the closure of the line, and second one of bridging the contour to complete it.
The initialization of the snake is a very important task, because the success of the snake adaptation to the image depends on it partly. To achieve the snake initialization over the image, it has been supposed that the hand has a similar shape than an ellipse. So it has been calculated the area of the binary image, its centre of masses and developed an open ellipse with equal orientation to the hand that binds it. So the initial snake is an open contour. Fig12, Fig13, and Fig14 show the adaptation of a snake to the contour of a hand, and the problem that can occur when the object is not completely in the image, as well as the improvement achieved with the open snake.

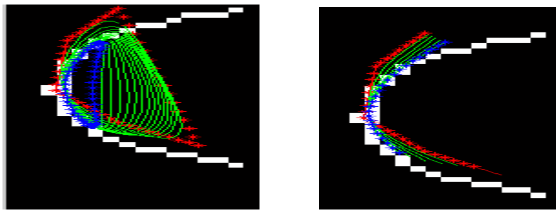

Fig.11

Fig.11 demonstrates how at the beginning the snake initialization is not close enough to the contour, but after several iterations, the snake fits the hand. However, when the fitting is not perfect the solution is to make more iteration, which increases the time computational cost. As Fig.12, Fig.13, and Fig.14 show the fitting is total in the second image of the sequence of images.

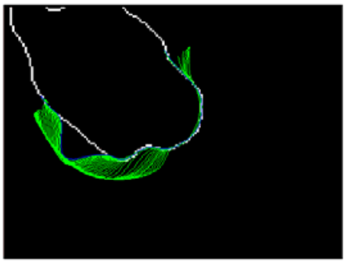

Fig.12

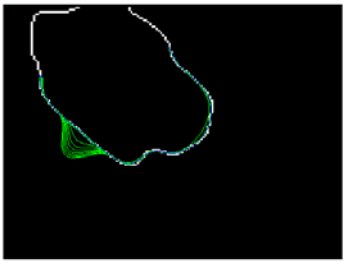

Fig.13

Fig.14

4 Practical applications of skin segmentation and active contours

Thus it was necessary to provide at the snake of a safety margin to reduce the risk of contact with the saw blade, and appoint the risk area.

To the first condition the final snake has been stretched adding a new energy to the snake energy defined.

$$E_{dilation}(v) = \lambda \frac{\sum_{i}^{n-1} v(i) - v(i+1)}{|v(s)|} \quad (17)$$

Where λ is a constant that can be positive or negative depending if it was the snake grows or diminishes, respectively.

$$E_{Snake}(v) = \frac{1}{2}\int_{\Omega} [E_{Int}(v) + E_{Ext}(v) + E_{dilation}(v)]\, ds \quad (18)$$

It has been established a region to prevent the snake from going into the risk zone, that is to say, into the area where the saw is located. In Fig.15 we can see this zone delimited by a blue line. The system compares this limit with the situation of the snake, when both meet a warning is arisen as it can be observed in Fig.16. In this way, this double security system would lower the risk factors in woodworking machinery.

Fig.15

WARNING!!!

Fig.16

5 Conclusions

A comparative study has been carried out between different colour regions in order to get skin segmentation in a wooden background mostly. The findings suggest that YCbCr space colour applied to a Bayes filter would be appropriate. And it has been found that wood characteristics are not so similar to the skin ones. Tracking the hand movements has been implemented with open active contour that fits their movements, and let know the position the hand has. Moreover a energy of dilation has been added to the snake that makes it stretched around the hand.

If illumination was dramatically modified there would be errors in the segmentation of the image particularly when working with a Bayes classifier. In future an advisable improvement o a neuronal network able to attain a much greater discrimination between two classes, skin and non-skin ones, even in the most critical conditions of lighting changes.

References:

[1] Blake, A., Isard, M., 1998. Active Contours. Edited by Springer-Verlag.

[2] Criado L., 2002. Identificación de la papila usando el algoritmo de contornos activos. Departamento de Inteligencia Artificial (UNED)

[3] Jure Kovac. 2003. Human Skin Clustering for Face Detection, EUROCON2003-International Conference on Computer as a Tool.

[4] Leymarie, F.F., 1990. Tracking and describing deformable objects using active contour models. McGill University

[5] Vladimir Vezhnevets, 2003, A Survey on Pixel-Based Skin colour Detection Techniques, Proc. Graphicon- pp85-92

[6] Xu, C. 1998. Snakes, Shapes, and Gradient Vector Flow. IEEE Transactions on Image Processing, vol.7, no 3. pp 359-369

[7] Xu, C., Pham D. L., Prince, J.L., 2000. Medical Image Segmentation Using Deformable Models. Vol2: pp. 129-174, edited by J.M. Fitzpatrick and M. Sonka, SPIE Press