



HAL
open science

Determining the Best Colour Axis for Detecting and Extracting Segments in Vision Robotics

Christophe Montagne, Camille Edie Nzi, Sylvie Lelandais, Jean Triboulet,
Humberto Loaiza

► **To cite this version:**

Christophe Montagne, Camille Edie Nzi, Sylvie Lelandais, Jean Triboulet, Humberto Loaiza. Determining the Best Colour Axis for Detecting and Extracting Segments in Vision Robotics. *Journal of Applied Science and Technology (JAST)*, 2006, 11, pp.5–14. hal-00342243

HAL Id: hal-00342243

<https://hal.science/hal-00342243>

Submitted on 1 Feb 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Determining the best color axis for detecting and extracting segments in vision robotics

C. MONTAGNE¹, C. N'ZI², S. LELANDAIS¹, J. TRIBOULET³, H. LOAIZA⁴

¹IBISC - CNRS FRE
2873
Université d'Evry
40, rue du Pelvoux
91020 Evry Cedex
France

²INPHB Yamoussoukro
(dpt Génie Electrique)
BP 1093 Yamoussoukro
Côte d'Ivoire
nziedie@yahoo.fr

³LIRMM - CNRS UMR 5506
Université de Montpellier 2
161, rue Ada - 34392
Montpellier Cedex 5
France

⁴School of Electrical
and Electronic
Engineering
University of Valle –
Santiago de Cali
COLUMBIA

ABSTRACT:

Our laboratory develops a system of assistance with a mobile robot to help disabled people in their current life. The environment is characterized by structured elements like segments.

In order to permit the localization and the navigation of a mobile robot within an interior environment, we have built a stereoscopic sensor and implemented all the algorithms which allow to obtain 3D coordinates of real objects from data images. Processing on the images leads us to have segments on the two images.

These segments are obtained from two images issued of a color stereoscopic sensor put on a mobile robot. The matching technique uses two classifiers. We calculate two probabilities to decide if this pairing is suitable. One probability is obtained with a neural classifier; the other is given by a bayesian method. To compute the probability, the two classifiers use the difference between lots of features calculated on the two segments. A previous work, performing on grey level images, has shown the interest of this approach and led us to select eight parameters. Now we work with color cameras, so we have to use this information to increase the good matching percentage.

In this paper, we present two methods in order to determine the best color axis. The first method is based on detected segments. In the second method, we extract a sample. By the use of the sample image we will perform the computation of color features to increase the segmentation.

1. INTRODUCTION

The LSC(Laboratoire Systèmes Complexes) is implied in the robotics of assistance to the disabled people through ARPH(Assistant Robot for Handicapped People) project [1]. Developed with the AFM (French Association against Myopathy) support, this project consists in developing a mobile robot supporting one camera and one arm manipulator, and guided remotely by the handicapped person. The operator can lead the robot manually, by direct or indirect vision (via a camera), or the robot moves automatically towards the place or the object indicated by the operator. The environment is characterized by a great diversity of colored and textured "objects". Thus, it is interesting to exploit color and texture information [2] in localization and navigation process. Moreover, greater part of the previous objects is structured and it can be assimilated to segments. Consequently, it is recommended to use segment primitives in the same processes. In a previous work [3] [4], we have shown it is possible to locate and to guide a base with a stereoscopic vision in gray levels. The segment building process was based on the DERICHE gradient [5] and on hysteresis segmentation. By matching the segments from two stereoscopic images, the 3D representation of the scene was built. Now, we use color cameras on the robot because this information is of primary importance for any indoor robotics application. So, we have to integrate color information, which is more complete than luminosity information, in the vision process. There are many manners to represent the colors.

In this paper, we present two methods in order to determine the best colour axis. The first method is based on detected segments. We apply algorithms to each colour axis or space and we analyze

detected segments with usual statistics like segment quantity, mean size and standard deviation. In the second method, we use the “Baker’s Transform” [6] to extract a sample, then we determinate the best colour axis if we have similarity between original and sample images.

We have decided to limit the number of spaces in our work. We have chosen five spaces which are representative of the whole of spaces. First, RGB(Red, Green and Blue)space is the most common space. It is habitual and simple. It is also the default space for bitmap images we use. Secondly, xyz space [5] is normalized and gives the chromaticity diagram. Thirdly, $L^*a^*b^*$ space [7] is a space of antagonistic axes which approaches the human visual system. L^* is axis of opposition between black and white, a^* between green and red, b^* between blue and yellow. Fourthly, HSI space [8] is an intuitive space, which estimates colors are composed of a hue, saturation and intensity. Fifthly, $I_1I_2I_3$ space [9] is based on principal components analysis. I_1 is axis of intensity, I_2 is axis of opposition between red end blue, and I_3 is axis of opposition between green and the two others.

2. FIRST METHOD: DETECTED SEGMENTS

We consider three different approaches to obtain segment sets. These approaches are different in their way to integrate color information in the image segmentation process. An indoor environment is characterized by a great number of structured objects. These objects are polyhedral or cylindrical. Many lines are visible which are verticals, horizontals or perspectives. By definition, segments are greatly present in an indoor scene. Moreover, one of the process objectives is to build a 3D representation, which requires use of segments.

2.1. Approach 1: Marginal segmentation

This approach is an application of the previous method [3], which was developed for gray level images. As it is shown on figure 1, from each color axis image ($i=1$ to 3), edge points of the scene objects are extracted. We use DERICHE gradient method [5] to obtain cartesian components of the gradient (horizontal and vertical components). Then usual method permits to have polar components of the gradient (magnitude and angle). Then images are binarised from hysteresis thresholding. Finally, segments are created from hysteresis segmentation.

2.2. Approach 2: Vectorial segmentation

Previously, we build segments from each color axis image. Now, we have to build segment by combination of the three color component images. As it is shown on figure 1, from each axis color image, we extract horizontal and vertical components of the gradient with DERICHE method, as previously. But we use DI ZENZO gradient [10] to obtain gradient magnitude and gradient angle images because it is a “component helping” method. Compared to the marginal process, from three inputs, we obtain a single gradient magnitude image and a single gradient angle image. Then, the process is the same. Finally, our output result is a single set of segments.

2.3. Approach 3: Edge “fusion” segmentation

As it is shown on figure 1, in this way, we have decided to conserve the process until having six edge binary images (three high binary images and three low binary images). Then, we create a single high binary image and a single low binary image by logical operation “OR”. Then, process continues as previously.

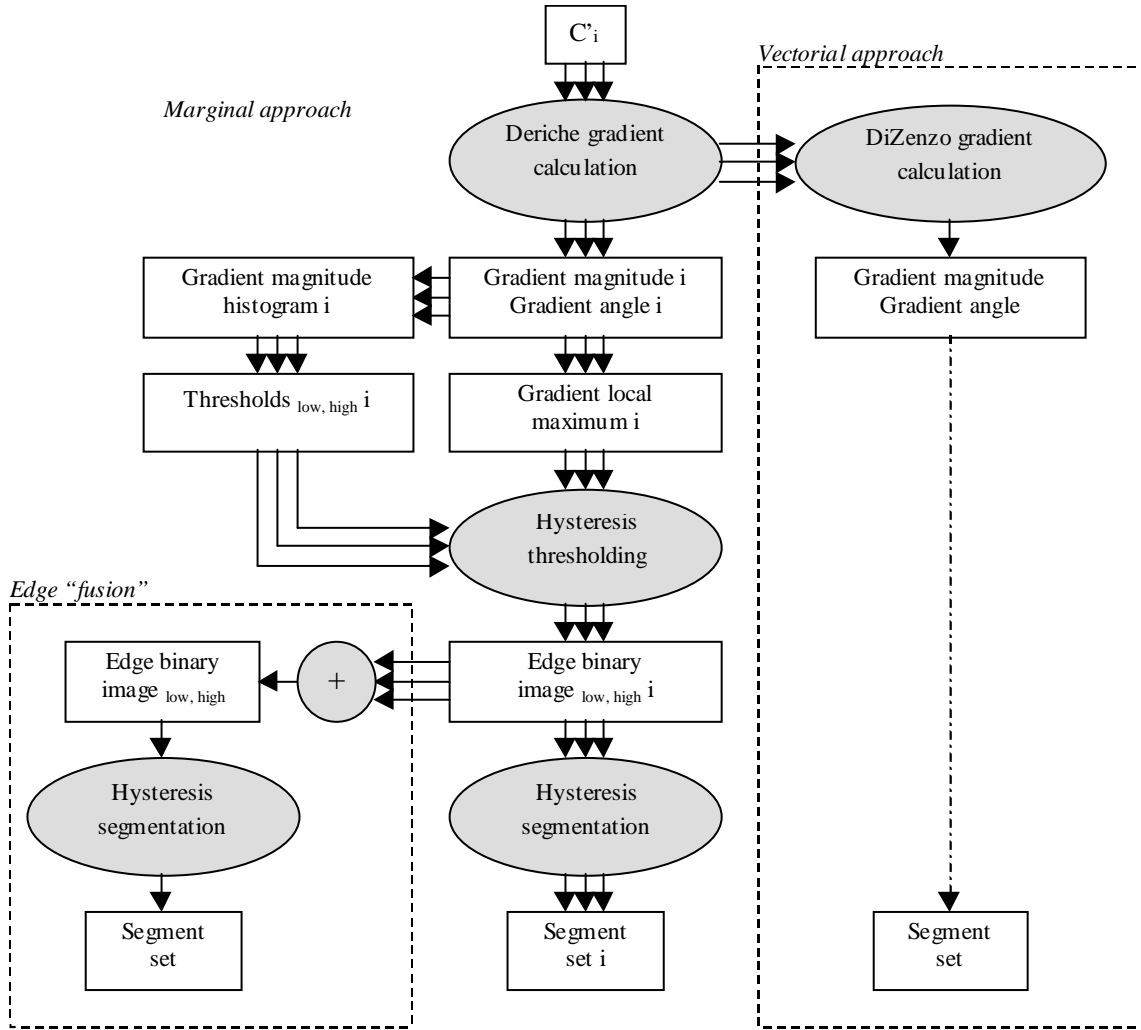


Figure 1: Three segment building approaches.

2.4. Notation

The quality of localization and navigation depends on the quality of the matching step results. These results will be better if extracted segments are correct (i.e. a sufficient number for each image, good value of the descriptors, ...).

To choose an adapted color space, we need measuring tools giving useful comparative data. Because few works have been done about this particular subject (some works concern specialized applications and can not be directly used [11][12]), we have developed our own tools. Detected segments are a fundamental criterion. For robot localization and navigation, segments must be long. Indeed, we suppose a long segment corresponds to a significant scene element. Moreover, during localization task, it is possible to forget small elements like plant foliage, ground squares, texture elements, and decorative objects.

We apply algorithms to each color axis or space and we analyze detected segments with usual statistics like segment quantity, mean size and standard deviation (of all segments or those having minimal size). We use also a notation which can be absolute ("weighted notation") or relative ("segments histogram comparison") as defined below.

2.4.1. Weighted notation

We distribute built segments on a size scale whose step (in pixels) is fixed arbitrarily. Then, for each step, we calculate a notation by multiplying the step index i with the quantity q of associated segments. The final note is the sum of all notes (eq. 3.1).

$$note = \sum_{i=0}^n (i \cdot q_i) \quad (eq1)$$

This computation favors great segments. So, the color axis or space, which provides the larger number of great segments, will obtain the best note and will be considered as the better axis. We have chosen five most significant color spaces (RGB, xyz, HSI, L*a*b*, I₁I₂I₃), so we have fifteen color axes. On all the images from database, we have to apply this method and compare the obtained notation in order to select three axes from different spaces, or one complete space. In some case, the step is equal to 25 pixels. So segments which have a size less than 25 pixels have a null contribution.

2.4.2. Segments histogram comparison

In order to evaluate the quality of colored image with regard to black and white image from the same scene, we compare histograms of the detected segments from these two images. The gray level image is obtained with the mean of the three components of the RGB space. This corresponds to the I₁-axis. With these two histograms, we compute also a notation, as explained below:

$$note = \sum_{i=0}^n (\Delta q_i \cdot i)$$

$$\Delta q_i = \sum_{j=low_i}^{j=high_i} (q - q_r)_j \quad (eq2)$$

With i the step index as previously defined, q the segment quantity from the colored image, and q_r the segment quantity from gray levels image reference. For a given step i , we have a minimal length (low_i) and a maximal length ($high_i$) of segments.

Note that this last approach includes the weighted notation, but it allows comparing each color axes with the same reference.

2.5. Outputs

For each approach, we present a result on figures 2 to 5 from RGB space. Note displayed segments have a minimal length of 20 pixels. In table 1, we present the number of segments and the mean length of the segments for the three approaches. With this example, we note a space does not give systematically more segments than an axis. In the same way, to have more segments does not indicate segments are larger. Thus, it is necessary to have tools to obtain information more significant. It is also necessary to cross these data.

Table 1: Some statistics examples.

	B - approach 1	RGB - approach 2	RGB - approach 3
Number of segments	658	383	1114
Mean length (pixels)	20	28	17

We compute different values from the image database but three are significant: quantity of segments, average size of segments, and segment histogram comparison. We compare at the same time the three segmentation approaches and the five color spaces, as it is shown on table 2.

We see approach 3 (edge “fusion”) gives more segments than the two others. Approach 2 based on the Di Zenzo gradient gives fewer segments. We show approach 2 gives larger segments and approach 3 smaller segments. Explanation is that approach 3 is based on a single gradient, which is optimized. Thus, segments are fewer but better. On the other hand, approach 2 gives edges more complete at the segmentation algorithm. Thus more segments are created and most of them correspond of scene details. The relative notation is significant.

This measure gives two types of information. First one is a relation between quantity and length of segments. So, we see approach 2 is worst because it does not provide enough segments. Many details are lost. Approach 3 is the best because object details are perceived better. Last one is the comparison between color spaces and gray levels. In this way, RGB space looks better than L*a*b*, I₁I₂I₃ and HSI. The xyz space seems to be less interesting. Moreover, except for RGB space with approach 3, color spaces are not better than intensity axis. Spaces with one intensity component have bad results because the two others components, which support color information, deteriorate the intensity information. We can say intensity axes are the best in order to extract segments. RGB space gives good results because it is a space where the intensity is supported by the three axes.



Figure 2: Original image of a chair.

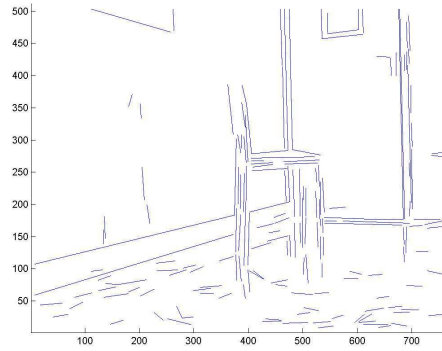


Figure 3: Segments from blue axis (approach 1).

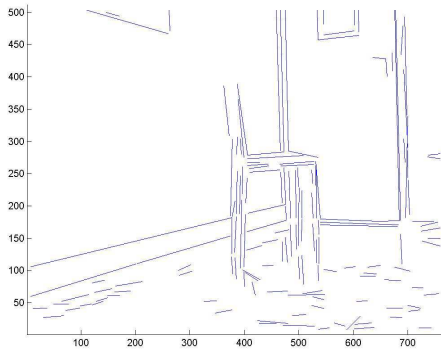


Figure 4: Segments from rgb space (approach 2).

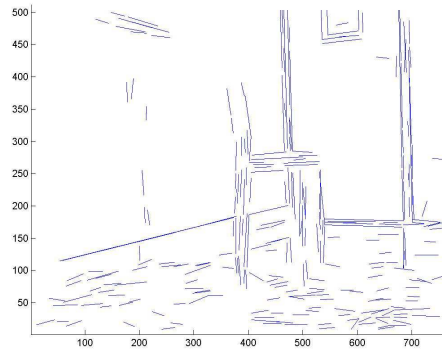


Figure 5: Segments from rgb space (approach 3).

Table 2: Some results from the different approaches to extract segments.

Average segment quantities					
	RGB	xyz	HSI	L*a*b*	I1I2I3
Approach 1	613	826	879	806	421
Approach 2	499	794	475	537	256
Approach 3	1664	1398	1387	1390	1102
Average segment lengths (in pixels)					
	RGB	xyz	HSI	L*a*b*	I1I2I3
Approach 1	32	22	23	24	30
Approach 2	36	24	33	37	59
Approach 3	18	17	18	18	21
Segments histogram comparison					
	RGB	xyz	HSI	L*a*b*	I1I2I3
Approach 1	-45	-184	-147	-120	-130
Approach 2	-66	-176	-230	-102	-183
Approach 3	17	-181	-123	-95	-74

3. SECOND METHOD: USING SAMPLE OF THE IMAGE

Now we have samples from original images, we project them on different color axes. These axes are issued from some varied color spaces. Note use of sample is very interesting in order to reduce computation time. To project one sample of 256 pixels on about twenty color axes is faster to project one image of full size on the same axes.

We use standard deviation to estimate color dynamic.

3.1. The Baker's Transform

On each image (of size 256x256 for example, but that can be more), we use the “Baker's Transform” to extract a sample of size 16x16. The Baker's Transform (BT for short) is based on the definition of mixing dynamical systems [13]. The main interest of these transformations is they mix in a very homogeneous way all the elements of the involved space.

One iteration of the BT is based on two steps: first, use an “affine” transformation, which gives an image twice larger and half high (the total number of pixels remains unchanged). Then, cut vertically in the middle the resulting image and to put the right half on the left half (the final image has the same size as the original image). The BT requires that the image size is $2^n \times 2^n$ pixels and we can show that the BT is periodic with period equal to $4N$ iterations. The image is well mixed with N iterations as shown in figure 6 (obtained from image on “figure 1a”). If we divide the mixed image and take a $2^p \times 2^p$ resulting window, we can obtain a good version of the original image at a reduced scale after applying $3p$ iterations of the BT to the mixed $2^p \times 2^p$ window (figure 7).

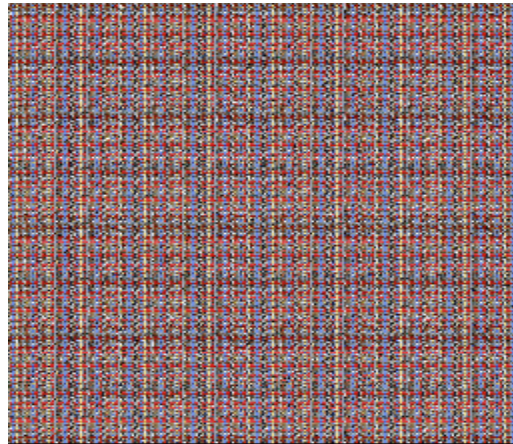


Figure 6: Transformed image at the 8th iteration.

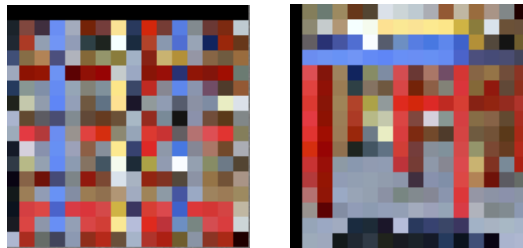


Figure 7: $2^4 \times 2^4$ windows from the mixed image and result after 12 iterations (after zoom).

In order to verify these samples are representatives of initial images, we use database, which contains more of 300 color images from different home rooms (kitchen, bathroom, living room) with varied decorations (modern, colored, wood). We calculate the mean of each color axis (R, G and B) and the mean of grey levels. We perform the same computation on all the original images and on all the 16x16 samples obtained from these images by the “Baker's Transform”.

figure 8 and figure 9 shows these values. We can see the similarity of the curves between the two figures. Tables 3 and 4 present mean values and standard deviation computed from all images. These tables confirm previous remarks about the similarity between original and sample images. Consequently, we can use sample to perform the choice of color axis, which is the best to characterize areas along segments. If, as in our example, the original image is of size 256x256 then this method enables us to consider 256 pixels instead of 65.536 to do our choice. The reduction of the computing time is undeniable because we divide by 256 the number of pixels used in the computation.

Table 3: Mean values of each component.

Image	Red	Green	Blue	Grey
Original	132,41	119,79	104,85	119,02
Sample	131,06	118,77	104,20	118,02

Table 4: Standard deviation of each component.

Image	Red	Green	Blue	Grey
Original	28,68	26,65	34,45	27,92
Sample	29,61	27,96	35,47	29,23

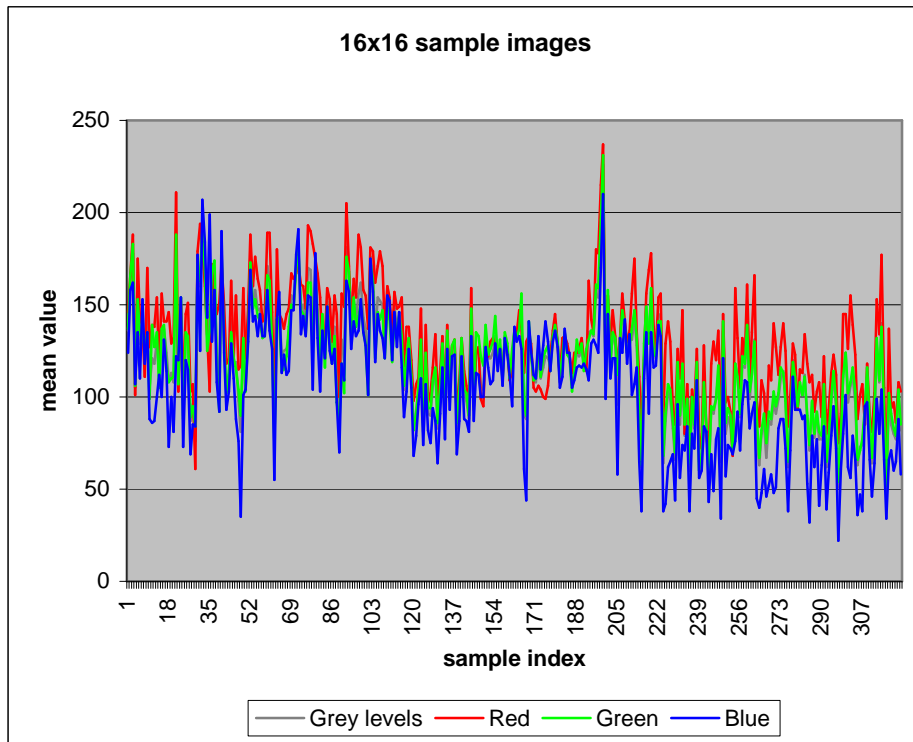


Figure 8: Mean values of each component from original images.

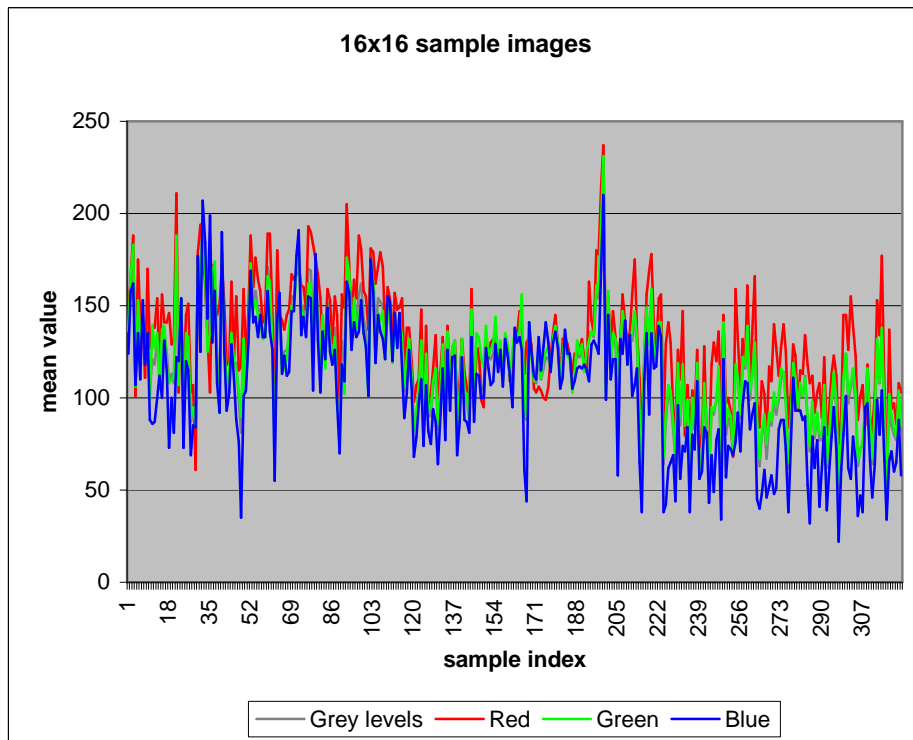


Figure 9: Mean values of each component from sample images.

3.2. Outputs

Now we have samples from original images, we project them on different color axes. These axes are issued from some varied color spaces. Note use of sample is very interesting in order to reduce computation time. To project one sample of 256 pixels on about twenty color axes is faster to project one image of full size on the same axes.

We use standard deviation to estimate color dynamic. If standard deviation is great, it means contrast will be important between different areas of the image. So color difference between two areas along segment will be discriminate. Table 5 shows, for some axes, the number of samples, which gives the higher contrasts. This table is obtained from the following way. For one sample, we compute standard deviation for all axes. Then, we look for the maximum of these values. Finally, all axis scores are incremented by one if their standard deviation is superior to 90% of the previous maximum value.

In Table 5, we can see that. H-axis was not retained because its cyclic nature is not adapted to a standard deviation calculation.

Table 5: Score of tested axes.

	R	G	B	x	y	z	H	S	I
Score	145	119	156	18	9	25	-	57	93
Axis	L*	a*	b*	I1	I2	I3	X	Y	Z
Score	125	6	23	93	0	0	203	204	220

4. RESULTS

Les résultats sur certainement la détection et l'extraction de segments. S'il n'y a pas de nouveaux résultats, on peut reconduire les résultats dans le cas des images Noir et Blanc (niveaux de gris) que nous avons obtenus et indiquer que les résultats dans le cas de l'image couleur sont en étude (voir notre article).

5. IMPORTANCE OF RESULTS FOR HANDICAPED PERSONS (relever l'importance des résultats dans l'aide des handicapés)

6. CONCLUSION (à compléter certainement)

This paper talks about the use of color information to increase the good matching percentage of segments. These segments are obtained starting from two images taken by a color stereoscopic sensor. In a first method, we have explained how it is possible to choose the best color axis based on detected segment. The results show that the five color axes used are not bad. The number of detected segments is significant. Therefore, the xyz space seems to be less interesting. Moreover the comparison between color spaces and gray levels show that the RGB space looks better than L*a*b*, I₁I₂I₃ and HIS. In a second method, we have explained how it is possible to choose the best color axis to perform the computation of color features. We extract 16x16 samples from original image by the use of the Baker's Transform, and then we select the color axis which gives the maximum standard deviation value. The table4 show that the RGB color axes and the L-axis give good result (score of R=145, G=119, B=156 and L=125) compare to the other axes. The two methods show that color axes R, G, B and L are interesting to exploit color and texture information in localization and navigation process.

Today, choice of color axis is operational. The next work is the correspondence of Line segments between two color images. We have to match them in pair, one from the top camera and the other from the bottom camera [14]. We will use Bayesian and neuronal approaches in comparison to the epipolar one [15]. By the use of the Baker's Transform we will perform the computation of color features to increase the segmentation. We will use two types of color parameters. One is a distance between colors in areas close to the segment, and the other one is an estimate of the color homogeneity in the same areas. After the matching, we perform the 3D reconstruction.

REFERENCES

1. P. Hoppenot, E. Colle, O. Ait Aider, Y. Rybarczyk, "ARPH - Assistant Robot for Handicapped People - A pluridisciplinary project", *IEEE Roman*, France, 2001, 624-629.
2. C. Montagne, S. Lelandais, J. Triboulet, C. N'Zi, "How to choose the best color space for the guidance of an indoor robot ?", CGIV 2002 - The First European Conference on Colour in Graphics, Image and Vision, Poitiers, France, April 2-5, 2002, pp. 589-593.
3. H. Loaiza, J. Triboulet, S. Lelandais, C. Barat, "Matching segments in stereoscopic vision", *IEEE Instrumentation & Measurement Magazine*, vol.4, n°1, 2001, 37-42.
4. H. Loaiza, «Capteur de vision stéréoscopique multiconfiguration pour la localisation d'un robot mobile», Robotic Thesis, University of Evry Val d'Essonne, France, November 18 1999.
5. R. Deriche, "Using Canny's criteria to derive an optimal edge detector recursively implemented", *The International Journal of Computer Vision*, 1987, 167-187.
6. Ph. Cornu, A. Smolarz, «Caractérisation d'images par textures associées», *Traitement du Signal*, vol.19, 2002, n°1.
7. A. Tremeau, "Représentation numérique des images couleur", *Spring School – Color Images*, Pau, France, March 2001.
8. R. C. Gonzales, R. E. Woods, "Digital Image Processing", *Addison-Wesley Publishing Co.*, Massachusetts, 1993, 221-251.
9. Y. I. Ohta, T. Kanade, T. Sakai, "Color information for region segmentation", *Computer Graphics and Image Processing* 13, 1980, 222-241.
10. S. Di Zenzo, "A note on the gradient of a multi-image", *Computer Vision, Graphics and Image Processing*, vol.33, 1986, 116-125.
11. J. C. Terrillon, Y. Niwa, K. Yamamoto, "On the selection of an efficient chrominance space for skin color-based image segmentation with an application to face detection", *QCAV*, Le Creusot, France, 2001, 409-414.
12. A. Albiol, L. Torres, E. J. Delp, "Optimum color spaces for skin detection", *International Conference on Image Processing*, October 2001, 122-124.
13. P. Billingsley, "Ergodic Theory and Information", John Wiley & Sons, Inc., New-York, 1965.
14. C. N'Zi, H. Loaiza, J. Triboulet, S. Lelandais, C. Barat, C. Montagne, "Matching stereoscopic segments: comparison between epipolar, bayesian and neuronal approaches", *ISMCR*, session B8, Bourges, France, 2002.
15. C. N'Zi, H. Loaiza, J. Triboulet, S. Lelandais, and C. Montagne, "Correspondance of line segments between two images: comparaison between epipolar, bayesian and neuronal approaches", *Global Journals of pure and applied sciences*, vol. 10, N°. 2, March 2004, pp. 335-341.