

A robust data association method based on B-Splines for SPLAM in complex environments

Alfredo Toriz P.[‡], Abraham Sánchez L., Rogelio Gonzalez V. and Beatriz Bernabe L.

[‡]Universidad Popular Autónoma del Estado de Puebla, Pue-México
Benemérita Universidad Autónoma de Puebla, Pue-México
`alfredo.toriz@upaep.mx`, `asanchez@cs.buap.mx`, `rgonzalez@cs.buap.mx`,
`bety@cs.buap.mx`

Abstract. In this work, we present a robust data association method for solving the SPLAM (Simultaneous Planning Localization and Mapping) problem in complex environments. The use of B-Splines curves as representation tool is one of the most important contributions. The problem of data association for SPLAM was initially addressed by comparing the control points forming the curves that represent the observed obstacles at a given moment, to the obstacles stored in the map, correlating those who are close enough. Then, the curvature of the related B-spline is obtained to extract characteristic points (inflection points and corners) contained in the curves. Finally, the matching information was used to correct the position of the robot and the detected obstacles. We carried out numerous experiments by using real and simulated information in order to validate the processes and algorithms proposed in our approach. Our method achieves a great precision in the map construction of complex environments, which is nearly impossible with techniques that currently exist.

1 Introduction

One of the fundamental challenges in robotics is to obtain robust and efficient mechanisms to model increasingly complex environments by using mobile robots or other kinds of robots. Many approaches have dealt with this problem without great success, since most of these methods try to extract geometric features and represent their positions on the map [1], or to discretize the space into cells [2] and then classify each as occupied or empty. These approaches are all limited to well-structured environments of limited size.

From these limitations, new methods have emerged in the literature that try to deal with this problem. One recent work uses B-Spline curves as a representation form for describing complex geometries with obstacles [3]. However, this new form of representation itself presents a new challenge in identifying similar environmental characteristics from successive scans and moments, so errors in the robot odometry can be corrected as the map of the environment is built. The approach presented by Pedraza et al., solves the problem of the representation

of complex geometries with obstacles by using a simple method of data association. This method involves projecting rays from a laser sensor from observed objects in the estimated odometry position to the located obstacles contained in the map, regardless of whether these points are really consistent. While this represents a real advance in the field, the method may lead to inconsistencies in the created map.

In this way, data association is undoubtedly the most critical aspect of the algorithm for map construction and it is considered by some authors to be the most difficult problem in this field of robotics. Therefore, the correct association between the sensed characteristics during an observation and the generated map features are essential for a consistent map construction, since any false association could invalidate the whole process.

The process of data association has received great attention in the robotics literature in recent years [4], [5] and many solutions have been proposed; However, most of these approaches are limited to simple features such as lines and points (preventing their use in complex environments), simple geometries, or simple curves where the association algorithms are very rudimentary. Although these techniques are appropriate and allow a full SPLAM in some cases, it is still necessary to develop and implement new strategies that may be applicable to more general cases without specific geometries.

In this paper, we present a new data association method for map construction of complex environments based on B-Spline curves, which are used to represent the obstacles in the map. Thus, our approach proposes an improvement on current methods of data association by fully exploiting the information contained in the parametric description of arbitrary geometries. With the creation and validation of an advanced map modeling method such as B-Spline curves, the properties of this representation (curves, length of curves, inflection points) are used to obtain a robust data association method applicable to most, if not all, mapping situations.

2 Data association based on B-Splines

2.1 Data management

When the robot gets a new set of measures of its environment through its perception system, the surrounding world is just a set of points $p_i \in \mathbb{R}^2$ linked to each other only by the logic provided by the scanning sensor (Figure 1). With these raw data, the first objective is to clearly identify the object associated with the measurements grouped into subsets to finally get the B-spline curves that represent the portions of the detected objects as closely as possible.

The segmentation of the raw data was performed with the adaptive cluster method proposed by Dietmayer [6] with the goal of obtaining measurements subgroups belonging to different objects. Once the subsets or clusters were obtained, the measurements corresponding to each detected object were approximated us-

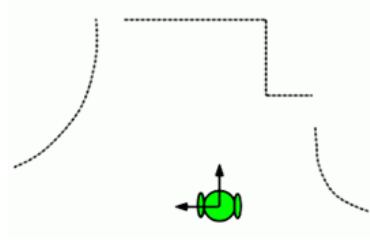


Fig. 1. Measurement points in space $p_i \in \mathbb{R}^2$ relative to an origin (the scanning robot) obtained with a laser scanning process.

ing unclamped B-spline curves of degree 3 as shown in the next equation:

$$D_k = C(t_k) = \sum_{i=0}^n N_{i,p}(t_k) X_i \text{ for } 0 \leq k \leq m. \quad (1)$$

where $\sum N_p(u)X$ is the Cox-de-Boor recursion formula to obtain a B-Spline curve, and the parameter t_k represents a point on the curve [7].

Although the representation based on B-Splines significantly reduced the noise caused by the measurement errors of the system; a smoothing of the curve was performed using a Gaussian filter to ensure that the process did not find false characteristic points. Then, an evolved version D_σ of the curve D was processed:

$$D_\sigma = \{x(u, \sigma), y(u, \sigma)\} \quad (2)$$

where

$$x(u, \sigma) = x(u) \otimes g(u, \sigma), \quad y(u, \sigma) = y(u) \otimes g(u, \sigma) \quad (3)$$

Here, $\otimes$ represents the convolution operator and $g(u, \sigma)$ denotes a Gaussian filter of amplitude σ , chosen conveniently and used only to remove the noise. While the valuable information about the curve was preserved, this selection process made the value to be very low. Since our proposed method normally works with open curves, a certain number of proportional points to the double of the full width at half-maximal (FWHM) of a Gaussian, were symmetrically offset on both ends of the curve when smoothing.

Finally, in order for the association process to effectively execute, it was necessary to ensure the invariance in the resolution of the curve. so, each discrete B-Spline curve was therefore stored taking equidistant points with a distance ε between each point:

$$\text{dist}(\sum N_p(u)X, \sum N_p(u+1)X) \cong \varepsilon \quad (4)$$

Thus, a parametric vector containing the B-Spline was obtained. In addition to the described process, a restriction on the length of the curve was applied; this due to the fact that too-small objects do not provide sufficient information and therefore it is not interesting to include them on the map. Also, although

our method is currently designed for use in static environments, this restriction allows filtering dynamic elements (people for example) that are not meant to be included in the map. Once the B-Splines were obtained and chosen, specific features contained in the curves were identified that were of great importance in the localization process. Essentially, two types of features were sought in the curves: inflection points and corners.

The corner type is very common feature representing a sharp point on a curve that is generally differentiable. On the other hand, the concept of inflection points in a very general way refers to the point on the curve where this passes from concave to convex or viceversa.

The process for obtaining both characteristics in the curve is based on the curvature scale space (CSS) [8] which is used to retrieve invariant geometric characteristics.

2.2 Search of characteristics in B-Splines curves

The search for particular characteristics in the curves that define objects is completely based on the curvature of the B-Spline, which is defined as the local measure that indicates how much a curve has moved away from a straight line. More formally, the curvature of a point $X_u = [x_u, y_u]$ on the B-Spline is defined as the amount equal to the inverse of the radius of the osculating circle at the point (the circle that touches tangentially to the curve in the X_u point), which means that the smaller is the radius of this circle ρ largest is the curvature at this point ($1/\rho$). The formula for processing the curvature can be expressed as:

$$k(u, \sigma) = \frac{\dot{x}(u, \sigma)\ddot{y}(u, \sigma) - \ddot{x}(u, \sigma)\dot{y}(u, \sigma)}{(\dot{x}(u, \sigma)^2 - \dot{y}(u, \sigma)^2)^{(3/2)}} \quad (5)$$

Where, in accordance with the convolution properties, the derivatives of each element could be easily calculated since the exact forms of the first and second derivatives of the used Gaussian kernels $\dot{g}(u, \sigma)$ and $\ddot{g}(u, \sigma)$ are known. Thus,

$$\dot{x}(u, \sigma) = \frac{\partial}{\partial u}(x(u) \otimes g(u, \sigma)) = x(u) \otimes \dot{g}(u, \sigma) \quad (6)$$

$$\ddot{x}(u, \sigma) = \frac{\partial^2}{\partial^2 u}(x(u) \otimes g(u, \sigma)) = x(u) \otimes \ddot{g}(u, \sigma) \quad (7)$$

$$\dot{y}(u, \sigma) = y(u) \otimes \dot{g}(u, \sigma) \quad (8)$$

$$\ddot{y}(u, \sigma) = y(u) \otimes \ddot{g}(u, \sigma) \quad (9)$$

Considering the above, the list of curvatures of the B-Spline is obtained by using the equation 5, i.e., the obtaining of the inflection points in the curve is carried out simply just searching for a change in the sign between two consecutive curvatures. Moreover, the searching of corners is based on the work of He et al., presented in [9]. Here, the detector is based on local and global properties of the curvature (Figure 2).

Since our proposed method uses open curves in most cases, we considered only corners delimited on both sides by inflection points as a safety and certainty method.

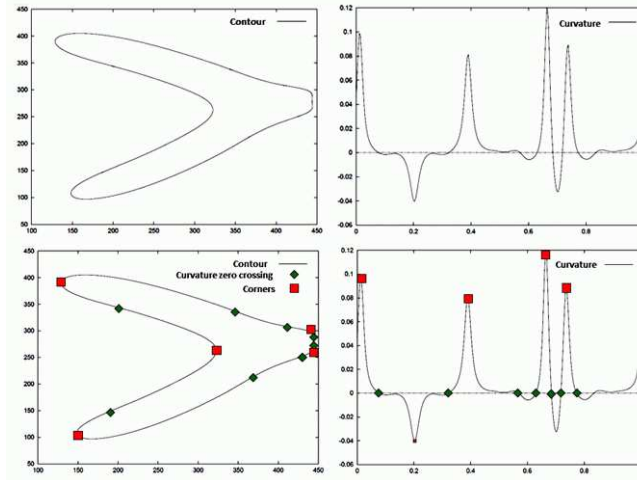


Fig. 2. Curvature zero crossings (inflections) and corners detection shown with some example curves to validate the detection and classification processes.

2.3 Data association

In order to understand the proposed approach, this section presents an illustrative example, which will be described in each stage of development. Figure 3 shows an example that corresponds to the evolution of the robot exploration until the instant $k + s$, where three objects $B_{O,1}$, $B_{O,2}$ and $B_{O,3}$ are detected (shown in red) in the sensor's range of vision (circumference in dotted line). This image highlights the odometric position as an a priori approximation of the robot's true position, obtained from the robot's motion model since errors and the sensory noise will lead the robot to a different real position and not to the expected one. For this reason, the objects $B_{O,1}$, $B_{O,2}$ and $B_{O,3}$ appear displaced and do not overlap exactly with the objects to which they belong.

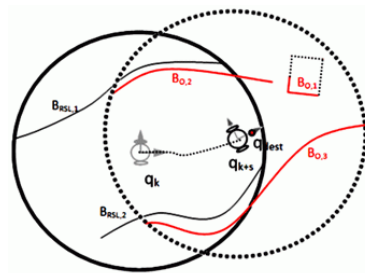


Fig. 3. The robot at the odometric position q_{k+s} with three obstacles detected within its detection range.

The B-Splines association will consider initially only the points of the control polygon that generated the curves. In this step, the distances between the points of the control polygons of the all objects (those contained in the Local Safe Region (see [10] for more details about the exploration method) and those observed in the position q_{k+s}) are obtained by associating the observed curves with the reference curves using the following criteria.

$$\min(dist(X_{LSR,i}, X_{O,j})) \leq d_{min}, \begin{cases} i = 1, \dots, n_m \\ j = 1, \dots, n_0 \end{cases} \quad (10)$$

Where $X_{O,j}$ and $X_{LSR,i}$ are the control points of the observed splines and the splines in the current LSR respectively, $dist(X_{LSR,i}, X_{O,j})$ represents the Euclidean distance between two points, and n_m and n_0 are the number of control points of the splines in the map and of the observed splines, respectively. At the end of this first stage, the splines with a minimum number μ_{min} of related control points will be associated.

In Figure 4, the curves $(B_{LSR,1}, B_{O,2})$ and $(B_{LSR,2}, B_{O,3})$ have been associated, since they have 5 and 7 related points of their control polygon respectively, which are obtained with equation 10.

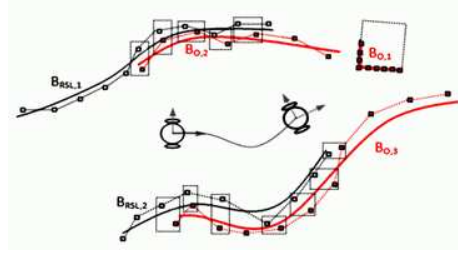


Fig. 4. Rough association computed with the control points of the curves.

If the curves are related at this point, the next step is to perform a fine association; for which, corners and inflection points contained in the related curves will be searched. If successful, the items found will be used for performing an accurate association between each pair of curves $(B_{LSR,i}, B_{O,j})$. The information about the type of characteristic and the curvature at that point will be used to avoid errors in this association step (Figure 5). If in some related curves no associated elements were found (lines or curves too smooth, for example), the fine association process will be executed to finding the closest points to the ends of the related curves as the authors proposed in [3].

Once all the elements are related, a search of the initial and final points of the related curves is performed. Taking the inflection points or the more extreme corners as starting points contained in both related curves, one can use some continuous points on the parametric curve to its end. The number of considered points is the maximum number of elements that can be taken in the curve

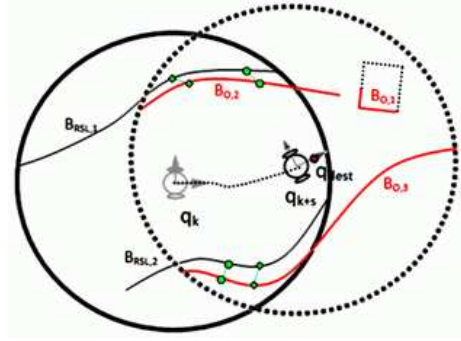


Fig. 5. Association of the inflection points between the curves and the corners of the local safe region and the observed curves. Green circles represent inflection points while diamonds represent corners.

segment of the smaller length of the two related points from the most extreme characteristic point toward the ends of the curve. This can be seen in detail in the related curves ($B_{LSR,2}, B_{O,3}$) in Figure 6. Here, one can note that the initial point, represented by the blue circle, was taken by choosing six elements of the parametric curve (they are showed in blue dashed line) from the inflection point, represented by the green circle, to the end of the curve. The length of the curve segment $B_{O,3}$ from the beginning of the curve to the inflection point, it is larger than the curve segment $B_{LSR,2}$ from the beginning of the curve to the inflection point and therefore the elements of the curve segment with shorter length will surely be contained in the another curve segment of greater length.

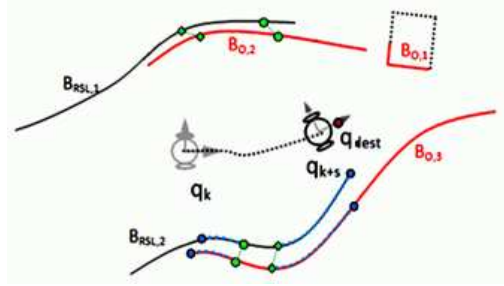


Fig. 6. Example of how the initial and the end point of the related curve segments are found.

The same process is performed at the end of the curve, where 15 points of the parametric curve were taken from the corner, represented with green diamond. The curve segment of smaller length from the corner at the end of the curve belonging to $B_{LSR,2}$ contains only 15 points. Finally, when the association

process ended, the related curves will have an appearance similar to those shown in Figure 7.

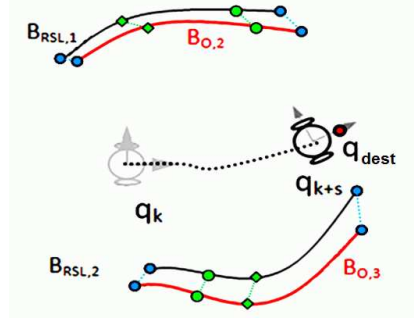


Fig. 7. The curve segments related to the described process of data association. Blue circles represent the ends of related curves, while green circles and green diamonds represent inflection points and the corners of the curve respectively.

3 Experimental results

Although the data association is a general problem that is applicable to any SPLAM algorithm, in this section we decided to implement our proposal by using the Extended Kalman Filter for SPLAM. This was done in order to make a comparison of our proposal with the proposal submitted by Pedraza et al. [3], since this is one of the few works that has grappled with the problem of complex environment representation using the same description of obstacles used in this work, though with a different approach to data association for our work. We have conducted numerous experiments with both real and simulated data in order to verify and validate the approach proposed in this work. The data obtained with simulation experiments have allowed the verification of the accuracy and consistency properties of the algorithm by comparisons with other existing and validated algorithms. Additionally, experiments with different data from real environments were used to verify the applicability and effectiveness of these techniques.

A simulated robot and the real Pioneer P3DX robot equipped with front and rear bumper arrays, a ring of eight forward ultrasonic transducer sensors (range-finding sonar) and a Hokuyo URG-04Lx laser range finder (with a detection range from 0.02 to 4 meters with a standard deviation (σ_L) of 1%, an angular resolution of 0.36 degrees and a scan angle of 240 degrees) were used in the experiments. The ultrasonic transducer sensors were used to obtain environmental information in the 120 degree range, since the laser range finder cannot see at this angle. The LIRMM laboratory environment was used in the experimental and simulation tests (the environment had several corridors), see Figure 8.

The data association strategy developed in this work and any other solution proposed in this field could be validated using a computational performance criterion, map quality and consistency of the algorithms. However, unlike the methods whose environment representation is based on specific geometries and where much of the information acquired by the sensors is wasted, our approach exploits the maximum amount of possible information provided by the sensors avoiding dangerous simplifications.

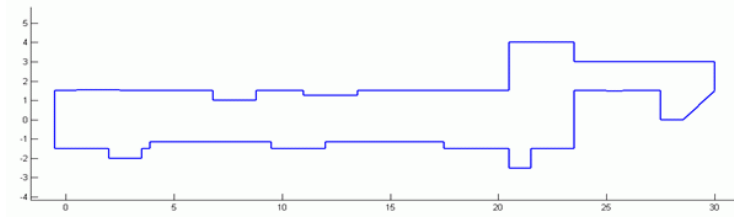


Fig. 8. The LIRMM environment.

Although the environment used in our experiments does not represent a serious challenge to the traditional algorithms, the use of a data association method based on B-Spline curves allows linking segments of continuous data, without the need for segmentation of information into smaller pieces. From the above, we find that the data association algorithm based on B-Splines, extends the capabilities of the SPLAM algorithms to more general situations. Figure 9 shows the results of SPLAM experiments we have developed as part of our research.

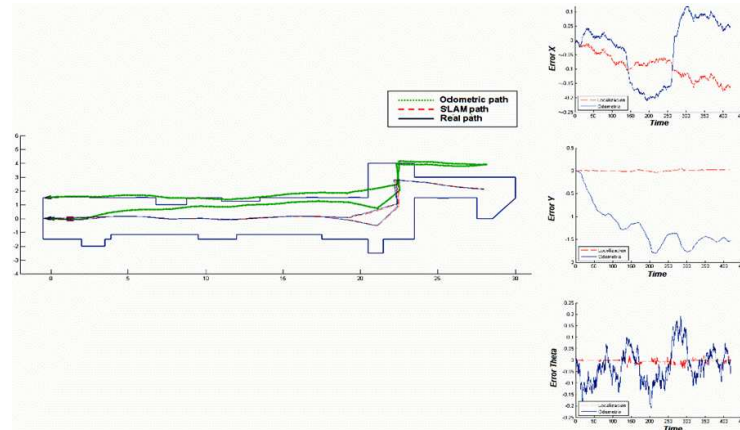


Fig. 9. Accuracy and consistency experiments for the EKF-SPLAM based on the data association approach proposed in this work.

Figure 9 shows the real trajectory of the robot (blue solid line), the odometric trajectory (green dotted line) and the trajectory obtained by the EKF-SPLAM method (dotted red line). On the other hand, when there is certainty about the real trajectory of the robot (as in the simulations) is possible to perform some checks to give an idea of the quality of the algorithms from the point of view of consistency. For this reason, it has been possible to include in the representation of the odometric error in X , Y and θ on the right part of the figure (blue line), as well as the localization errors (red dotted line). From the data obtained with our approach and taking as reference the errors shown by the SPLAM method presented by Pedraza et al. [3], we conclude that the proposed method of data association notably improve the quality of the map and the errors as shown in Figure 10.

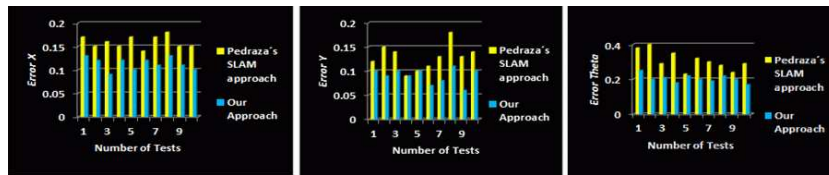


Fig. 10. Errors obtained with SPLAM strategies, for the simple data association [3] and for the proposed data association in this work.

Finally, in this section we present experiments with real data in real environments to validate the results presented in the simulation step. In all tests, the maps were obtained by considering only the odometric information reported by the robot and the maps obtained after applying our SPLAM implementation with our proposed data association. Figure 11 shows the real environment. The map obtained with our SPLAM strategy is shown in Figure 12a, while the odometric map obtained using only odometry information is shown in Figure 12b.



Fig. 11. The real environment used for the tests.

Considering that our approach has been proposed to build maps of complex environments, another experiment was performed using the real environment

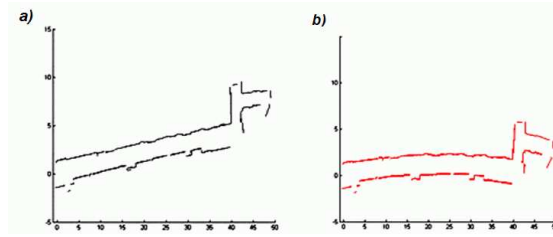


Fig. 12. a) Map built with the proposed SPLAM strategy. b) Map built with only odometry information.

shown in Figure 13. In this example, we can see that the construction of the map would be impossible using a conventional method based on lines and points. Figure 14 shows the map obtained with our approach.



Fig. 13. LIRMM environment with objects.

4 Conclusions

We have proposed a method for data association based on the curvatures analysis of the related B-Splines curves, for solving the SPLAM (Simultaneous Planning Localization and Mapping) problem in complex environments. Here, we use the techniques of digital images CSS [8], used in the pattern recognition field, inflection points and corners extraction among others. This association mechanism not only established a robust correspondence between the observations made by the robot and the objects contained in the map, but it facilitated parametric correspondence between each pair of associated representative elements.

After testing the proposed method in both simulated and real complex environments, and comparing the performance of the system to standard algorithms like EKF_SPLAM, we conclude that this functionality was well demonstrated and supports the strength of this new approach to this family of mapping problems.

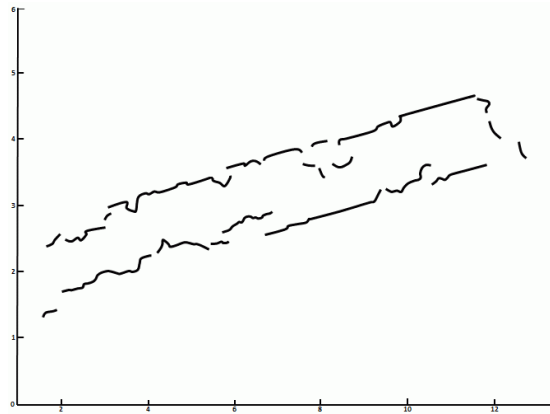


Fig. 14. Final map of the complex environment.

References

1. S. Zhang, L. Xie and M. Adams, "Gradient model based feature extraction for simultaneous localization and mapping in outdoor applications", *IEEE Control, Automation, Robotics and Vision Conference*, (2004) 431-436
2. G. Grisetti, C. Stachniss and W. Burgard, "Improving Grid-based SLAM with Rao-Blackwellized Particle Filters by Adaptive Proposals and Selective Resampling", *IEEE International conference on Robotics and Automation*, (2005) 2432-2437
3. L. Pedraza, D. Rodriguez-Losada, F. Matia, G. Dissanayake and J. V. Miro, "Extending the Limits of Feature-Based SLAM with B-Splines", *IEEE Transactions on Robotics*, Vol. 25, Issue 2, (2009) 353-366
4. A. J. Cooper, "A comparison of data association techniques for Simultaneous Localization and Mapping", *PhD Thesis, Massachusetts Institute of Technology - Dept. of Aeronautics and Astronautics*, (2005)
5. S. Thrun, M. Montemerlo, D. Koller, B. Wegbreit, J. Nieto and E. Nebot, "Fast-SLAM: An efficient solution to the simultaneous localization and mapping problem with unknown data association", *Journal of Machine Learning Research*, (2004)
6. K. C. J Dietmayer, J. Sparbert and D. Streller, "Model based object classification and object tracking in traffic scenes from range images", *Proc. of the IV IEEE Intelligent Vehicles Symposium*, (2001)
7. C. de Boor, "A practical guide to Splines", *Springer-Verlag, New York*, (1978)
8. F. Mokhtarian, "Silhouette-based isolated object recognition through curvature scale space", *IEEE Pattern Analysis and Machine Intelligence*, Vol. 17, No. 5, (1995) 539-544
9. X. C. He and N. H. C. Yung, "Corner detector based on global and local curvature properties", *Optical Engineering*, 47(5) (2008)
10. A. Toriz P., A. Sánchez L., R. Zapata and M. A. Osorio L., "Mobile robot SPLAM for robust navigation", *Journal Research in Computing Science*, Vol. 54, (2011) 295-306

Image Quantizer based on Contrast Band-Pass Filtering

Jaime Moreno^{1,2}, Oswaldo Morales¹, and Ricardo Tejeida¹

¹ National Polytechnic Institute of Mexico,
IPN Avenue, Lindavista, Mexico City, 07738, Mexico.

² XLIM Laboratory, Signal, Image and Communications Department,
University of Poitiers, 86962 Futuroscope, France.
jmorenoe@ipn.mx

Abstract The aim of this work is to explain how to apply perceptual criteria in order to define a perceptual forward and inverse quantizer. We present its application to the H_i-SET coder. Our approach consists in quantizing wavelet transform coefficients using some of the human visual system behavior properties. Taking in to account that noise is fatal to image compression performance, because it can be both annoying for the observer and consumes excessive bandwidth when the imagery is transmitted. Perceptual quantization reduces unperceivable details and thus improve both visual impression and transmission properties. The comparison between JPEG2000 coder and the combination of H_i-SET with the proposed perceptual quantizer (χ SET) shows that the latter is not favorable in PSNR than the former, but the recovered image is more compressed (less bit-rate) at the same or even better visual quality measured with well-know image quality metrics, such as MSSIM, UQI or VIF, for instance.

Keywords: Human Visual System, Contrast Sensitivity Function, Perceived Images, Wavelet Transform, Peak Signal-to-Noise Ratio, No-Reference Image Quality Assessment, JPEG2000.

1 Introduction

Digital image compression has been a research topic for many years and a number of image compression standards has been created for different applications. The JPEG2000 is intended to provide rate-distortion and subjective image quality performance superior to existing standards, as well as to supply functionality [1]. However, JPEG2000 does not provide the most relevant characteristics of the human visual system, since for removing information in order to compress the image mainly information theory criteria are applied. This information removal introduces artifacts to the image that are visible at high compression rates, because of many pixels with high perceptual significance have been discarded.

Hence, it is necessary an advanced model that removes information according to perceptual criteria, preserving the pixels with high perceptual relevance regardless of the numerical information. The Chromatic Induction Wavelet Model

presents some perceptual concepts that can be suitable for it. Both CBPF and JPEG2000 use wavelet transform. CBPF uses it in order to generate an approximation to how every pixel is perceived from a certain distance taking into account the value of its neighboring pixels. By contrast, JPEG2000 applies a perceptual criteria for all coefficients in a certain spatial frequency independently of the values of its surrounding ones. In other words, JPEG2000 performs a global transformation of wavelet coefficients, while CBPF performs a local one.

CBPF attenuates the details that the human visual system is not able to perceive, enhances those that are perceptually relevant and produces an approximation of the image that the brain visual cortex perceives. At long distances the lack of information does not produce the well-known compression artifacts, rather it is presented as a softened version, where the details with high perceptual value remain (for example, some edges).

2 JPEG2000 Global Visual Frequency Weighting

In JPEG2000, only one set of weights is chosen and applied to wavelet coefficients according to a particular viewing condition (100, 200 or 400 dpi's) with fixed visual weighting[1, Annex J.8]. This viewing condition may be truncated depending on the stages of embedding, in other words at low bit rates, the quality of the compressed image is poor and the detailed features of the image are not available since at a relatively large distance the low frequencies are perceptually more important. The table 1 specifies a set of weights which was designed for the luminance component based on the CSF value at the mid-frequency of each spatial frequency. The viewing distance is supposed to be 4000 pixels, corresponding to 10 inches for 400 dpi print or display. The weight for LL is not included in the table, because it is always 1. Levels 1, 2, $\dots$, 5 denote the spatial frequency levels in low to high frequency order with three spatial orientations, *horizontal*, *vertical* and *diagonal*.

Table 1. Recommended JPEG2000 frequency (s) weighting for 400 dpi's ($s = 1$ is the lowest frequency wavelet plane).

s	<i>horizontal</i>	<i>vertical</i>	<i>diagonal</i>
1	1	1	1
2	1	1	0.731 668
3	0.564 344	0.564 344	0.285 968
4	0.179 609	0.179 609	0.043 903
5	0.014 774	0.014 774	0.000 573

3 Perceptual Forward Quantization

3.1 Methodology

Quantization is the only cause that introduces distortion into a compression process. Since each transform sample at the perceptual image $\mathcal{I}_\rho$ is mapped independently to a corresponding step size either Δ_s or Δ_n , thus $\mathcal{I}_\rho$ is associated with a specific interval on the real line. Then, the perceptually quantized coefficients $\mathcal{Q}$, from a known viewing distance d , are calculated as follows:

$$\mathcal{Q} = \sum_{s=1}^n \sum_{o=v,h,d} \text{sign}(\omega_{s,o}) \left\lfloor \frac{|\alpha(\nu, r) \cdot \omega_{s,o}|}{\Delta_s} \right\rfloor + \left\lfloor \frac{c_n}{\Delta_n} \right\rfloor \quad (1)$$

Unlike the classical techniques of Visual Frequency Weighting (VFW) on JPEG2000, which apply one CSF weight per sub-band [1, Annex J.8], Perceptual Quantization using CBPF (ρ SQ) applies one CSF weight per coefficient over all wavelet planes $\omega_{s,o}$. In this section we only explain Forward Perceptual Quantization using CBPF (F- ρ SQ). Thus, Equation 1 introduces the perceptual criteria of Perceptual Images to each quantized coefficient of Equation of Dead-zone Scalar Quantizer. A normalized quantization step size $\Delta = 1/128$ is used, namely the range between the minimal and maximal values at $\mathcal{I}_\rho$ is divided into 128 intervals. Finally, the perceptually quantized coefficients are entropy coded, before forming the output code stream or bitstream.

3.2 Experimental Results applied to JPEG2000

The Perceptual quantizer F- ρ SQ in JPEG2000 is tested on all the color images of the *Miscellaneous volume* of the University of Southern California Image Data Base[2]. The data sets are eight 256×256 pixel images and eight 512×512 pixel images, but only visual results of the well-known images *Lena*, *F-16* and *Baboon* are depicted, which are 24-bit color images and 512×512 of resolution. The CBPF model is performed for a 19 inch monitor with 1280 pixels of horizontal resolution at 50 centimeters of viewing distance. The software used to obtain a JPEG2000 compression for the experiment is *JJ2000*[3]. Figure 1(a) shows the assessment results of the average performance of color image compression for each bit-plane using a Dead-zone Uniform Scalar Quantizer (SQ, function with heavy dots), and it also depicts the results obtained when applying F- ρ SQ(function with heavy stars). Using CBPF as a method of forward quantization, achieves better compression ratios than SQ with the same threshold, obtaining better results at the highest bit-planes, since CBPF reduces unperceivable features. Figure 1(b) shows the contribution of F- ρ SQ in the JPEG2000 compression ratio, for example, at the eighth bit-plane, CBPF reduces 1.2423 bits per pixel than the bit rate obtained by SQ, namely in a 512×512 pixel color image, CBPF estimates that 39.75KB of information is perceptually irrelevant at 50 centimeters.

Figure 2 depicts examples of recovered images compressed at 0.9 bits per pixel by means of JPEG2000 (a) without and (b) with F- ρ SQ. Also these figures

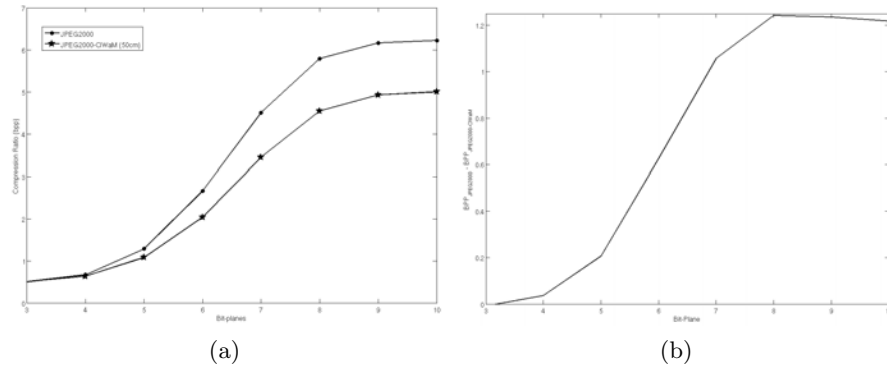


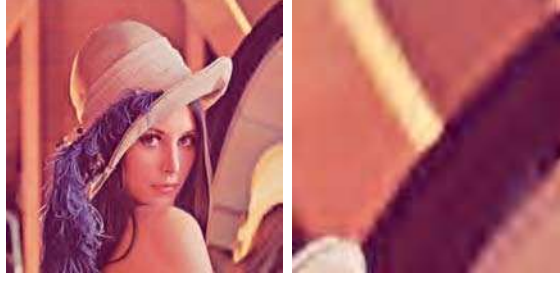
Figure 1. (a) JPEG2000 Compression ratio (bpp) as a function of Bit-plane. Function with heavy dots shows JPEG2000 only quantized by the dead-zone uniform scalar manner. While function with heavy stars shows JPEG2000 perceptually pre-quantized by F- ρ SQ. (b) The bit-rate decrease by each Bit-plane after applying F- ρ SQ on the JPEG2000 compression.

show that the perceptual quality of images forward quantized by ρ SQ is better than the objective one. Also, figure 3 shows examples of recovered images of *Baboon* compressed at 0.59, and 0.45 bits per pixel by means of JPEG2000 (a) without and (b) with F- ρ SQ. In Fig. 3(a) PSNR=26.18 dB and in Fig. 3(b) PSNR=26.15 dB but a perceptual metrics like WSNR [4], for example, assesses that it is equal to 34.08 dB. Therefore, the recovered image Forward quantized by ρ SQ is perceptually better than the one only quantized by a SQ. Since the latter produces more compression artifacts, the ρ SQ result at 0.45 bpp (Fig. 3(b)) contains less artifacts than SQ at 0.59 bpp. For example the *Baboon's* eye is softer and better defined using F- ρ SQ and it additionally saves 4.48 KB of information.

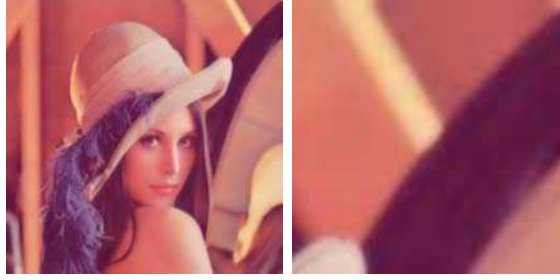
4 Perceptual Inverse Quantization

The proposed Perceptual Quantization is a generalized method, which can be applied to wavelet-transform-based image compression algorithms such as EZW, SPIHT, SPECK or JPEG2000. In this work, we introduce both forward (F- ρ SQ) and inverse perceptual quantization (I- ρ SQ) into the H $\hat{i}$ -SET coder. This process is shown in the green blocks of Fig. 4. An advantage of introducing ρ SQ is to maintain the embedded features not only of H $\hat{i}$ -SET algorithm but also of any wavelet-based image coder. Thus, we call CBPF Perceptual Quantization + H $\hat{i}$ -SET = cH $\hat{i}$ -SET or χ SET.

Both JPEG2000 and χ SET choose their VFWs according to a final viewing condition. When JPEG2000 modifies the quantization step size with a certain visual weight, it needs to explicitly specify the quantizer, which is not very suitable for embedded coding. While χ SET neither needs to store the visual



(a) JPEG2000 PSNR=31.19 dB.



(b) JPEG2000-F-ρSQ PSNR=27.57 dB.

Figure 2. Examples of recovered images of Lenna compressed at 0.9 bpp.

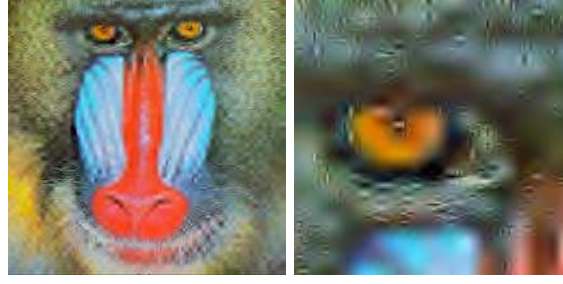
weights nor to necessarily specify a quantizer in order to keep its embedded coding properties.

The main challenge underlies in to recover not only a good approximation of coefficients $\mathcal{Q}$ but also the visual weight $\alpha(\nu, r)$ (Eq. 1) that weighted them. A recovered approximation $\hat{\mathcal{Q}}$ with a certain distortion Δ is decoded from the bitstream by the entropy decoding process. The VFWs were not encoded during the entropy encoding process, since it would increase the amount of stored data. A possible solution is to embed these weights $\alpha(\nu, r)$ into $\hat{\mathcal{Q}}$. Thus, our goal is to recover the $\alpha(\nu, r)$ weights only using the information from the bitstream, namely from the Forward quantized coefficients $\hat{\mathcal{Q}}$.

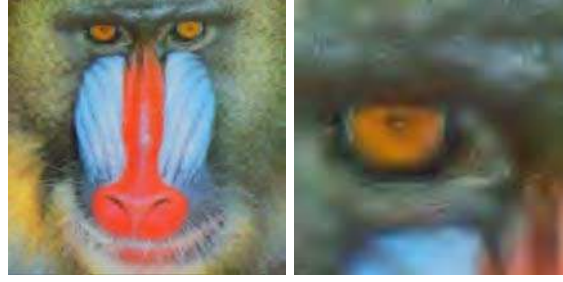
Therefore, our hypothesis is that an approximation $\hat{\alpha}(\nu, r)$ of $\alpha(\nu, r)$ can be recovered applying CBPF to $\hat{\mathcal{Q}}$, with the same viewing conditions used in $\mathcal{I}$. That is, $\hat{\alpha}(\nu, r)$ is the recovered e-CSF. Thus, the perceptual inverse quantizer or the recovered $\hat{\alpha}(\nu, r)$ introduces perceptual criteria to Inverse Scalar Quantizer and is given by:

$$\hat{\mathcal{I}} = \begin{cases} \sum_{s=1}^n \sum_{o=v,h,d} \text{sign}(\widehat{\omega_{s,o}}) \frac{\Delta_s \cdot (|\widehat{\omega_{s,o}}| + \delta)}{\hat{\alpha}(\nu, r)} + (\hat{c}_n + \delta) \cdot \Delta_n & |\widehat{\omega_{s,o}}| > 0 \\ 0, & \widehat{\omega_{s,o}} = 0 \end{cases} \quad (2)$$

For the sake of showing that the encoded VFWs are approximately equal to the decoded ones, that is $\alpha(\nu, r) \approx \hat{\alpha}(\nu, r)$, we perform two experiments.



(a) JPEG2000 compressed at 0.59 bpp.



(b) JPEG2000-F- ρ SQ compressed at 0.45 bpp.

Figure 3. Examples of recovered images of Baboon.



Figure 4. The χ SET image compression algorithm. Green blocks are the F- ρ SQ and I- ρ SQ procedures.

Experiment 1: Histogram of $\alpha(\nu, r)$ and $\hat{\alpha}(\nu, r)$. The process of this short experiment is shown by Figure 5. Figure 5(a) depicts the process for obtaining losslessly both Encoded and Decoded visual weights for the 512×512 *Lena* image, channel *Y* at 10 meters. While Figures 5(b) and 5(c) shows the frequency histograms of $\alpha(\nu, r)$ and $\hat{\alpha}(\nu, r)$, respectively. In both graphs, the horizontal axis represents the sort of VFW variations, whereas the vertical axis represents the number of repetitions in that particular VFW. The distribution in both histograms is similar and they have the same shape.

Experiment 2: Correlation analysis between $\alpha(\nu, r)$ and $\hat{\alpha}(\nu, r)$. We employ the process shown in Fig. 5(a) for all the images of the CMU, CSIQ, and IVC Image Databases. In order to obtain $\hat{\alpha}(\nu, r)$, we measure the lineal correlation between the original $\alpha(\nu, r)$ applied during the F- ρ SQ process and the recovered $\hat{\alpha}(\nu, r)$. Table 2 shows that there is a high similarity between the applied VFW and the recovered one, since their correlation is 0.9849, for gray-scale images, and 0.9840, for color images.

In this section, we only expose the results for the CMU image database.

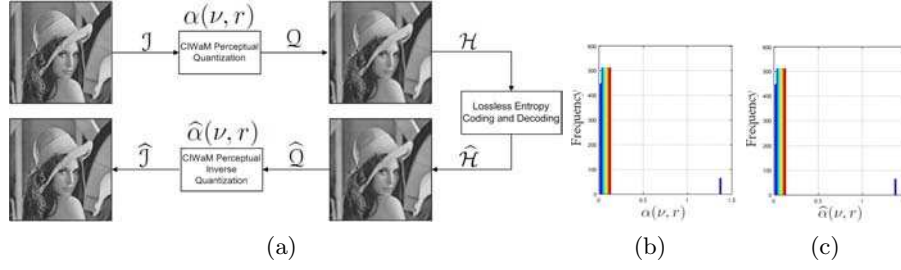


Figure 5. (a) Graphical representation of a whole process of compression and decompression. Histograms of (b) $\alpha(\nu, r)$ and (c) $\hat{\alpha}(\nu, r)$ visual frequency weights for the 512×512 image *Lenna*, channel *Y* at 10 meters.

Table 2. Correlation between $\alpha(\nu, r)$ and $\hat{\alpha}(\nu, r)$ across CMU, CSIQ, and IVC Image Databases.

Image Database	8 bpp gray-scale	24 bpp color
CMU	0.9840	0.9857
CSIQ	0.9857	0.9851
IVC	0.9840	0.9840
Overall	0.9849	0.9844

Fig. 6 depicts the PSNR difference (dB) of each color image of the CMU database, that is, the gain in dB of image quality after applying $\hat{\alpha}(\nu, r)$ at $d = 2000$ centimeters to the $\hat{Q}$ images. On average, this gain is about 15 dB. Visual examples of these results are shown by Fig. 7, where the right images are the original images, central images are perceptual quantized images after applying $\alpha(\nu, r)$ and left images are recovered images after applying $\hat{\alpha}(\nu, r)$. After applying $\hat{\alpha}(\nu, r)$, a visual inspection of these sixteen recovered images show a perceptually lossless quality. We perform the same experiment for gray-scale and color images with $d = 20, 40, 60, 80, 100, 200, 400, 800, 1000$ and 2000 centimeters, in addition to test their objective and subjective image quality by means of the PSNR and MSSIM metrics, respectively.

In Figs. 8 and 9, green functions denoted as F- ρ SQ are the quality metrics of perceptual quantized images after applying $\alpha(\nu, r)$, while blue functions denoted as I- ρ SQ are the quality metrics of recovered images after applying $\hat{\alpha}(\nu, r)$. Thus, either for gray-scale or color images, both PSNR and MSSIM estimations of the quantized image $\hat{Q}$ decrease regarding d , the longer d the greater the image quality decline. When the image decoder recovers $\hat{Q}$ and it is perceptually inverse quantized, the quality barely varies and is close to perceptually lossless, no matter the distance.

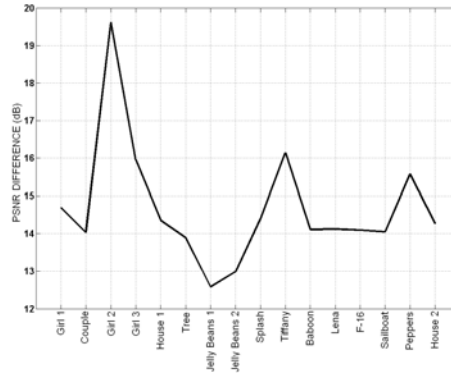


Figure 6. PSNR difference between $\hat{Q}$ image after applying $\alpha(\nu, r)$ and recovered $\hat{I}$ after applying $\hat{\alpha}(\nu, r)$ for every color image of the CMU database.

5 Conclusions

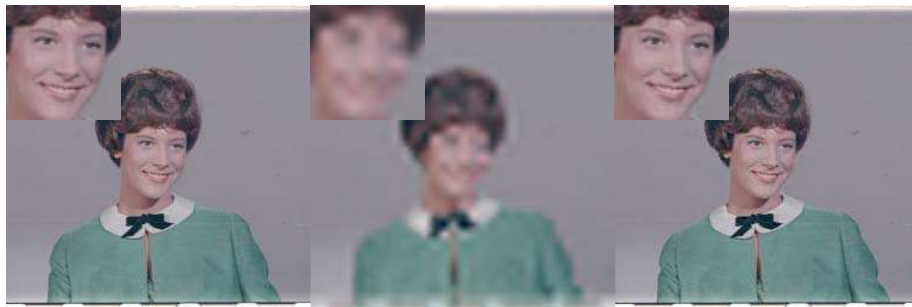
In this work, we defined both forward (F- ρ SQ) and inverse (I- ρ SQ) perceptual quantizer using CBPF. We incorporated it to Hi-SET, testing a perceptual image compression system χ SET. In order to measure the effectiveness of the perceptual quantization, a performance analysis is done using thirteen assessments such as PSNR, MSSIM, VIF, WSNR or $\mathcal{NR}$ PSNR, for instance, which measured the image quality between reconstructed and original images. The experimental results show that the solely usage of the Forward Perceptual Quantization improves the JPEG2000 compression and image perceptual quality. In addition, when both Forward and Inverse Quantization are applied into Hi-SET, it significantly improves the results regarding the JPEG2000 compression.

Acknowledgment

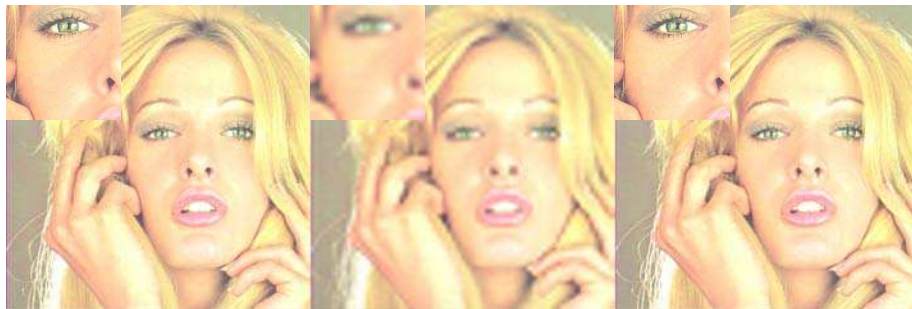
This work is supported by National Polytechnic Institute of Mexico by means of Project No. 20140096, the Academic Secretary and the Committee of Operation and Promotion of Academic Activities (COFAA), National Council of Science and Technology of Mexico by means of Project No. 204151/2013, and LABEX Σ -LIM France, Coimbra Group Scholarship Programme granted by University of Poitiers and Region of Poitou-Charentes, France.

References

1. M. Boliek, C. Christopoulos, and E. Majani, *Information Technology: JPEG2000 Image Coding System*, JPEG 2000 Part I final committee draft version 1.0 ed., ISO/IEC JTC1/SC29 WG1, JPEG 2000, April 2000.



(a) Girl 2



(b) Tiffany



(c) Peppers

Figure 7. Visual examples of Perceptual Quantization. Left images are the original images, central images are forward perceptual quantized images (F- ρ SQ) after applying $\alpha(\nu, r)$ at $d = 2000$ centimeters and right images are recovered I- ρ SQ images after applying $\hat{\alpha}(\nu, r)$.

2. S. I. P. I. of the University of Southern California. (1997) The USC-SIPI image database. Signal and Image Processing Institute of the University of Southern California. [Online]. Available: <http://sipi.usc.edu/database/>
3. C. Research, École Polytechnique Fédérale de Lausanne, and Ericsson. (2001) JJ2000 implementation in Java. Cannon Research, École Polytechnique Fédérale

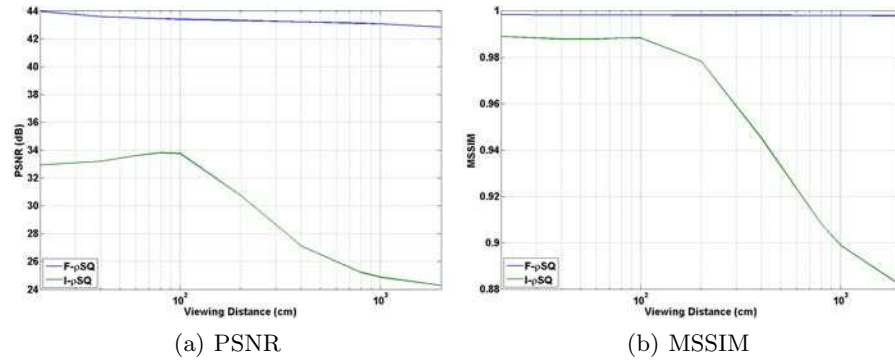


Figure 8. PSNR and MSSIM assessments of compression of Gray-scale Images (Y Channel) of the CMU image database. Green functions denoted as F- ρ SQ are the quality metrics of forward perceptual quantized images after applying $\alpha(\nu, r)$, while blue functions denoted as I- ρ SQ are the quality metrics of recovered images after applying $\hat{\alpha}(\nu, r)$.

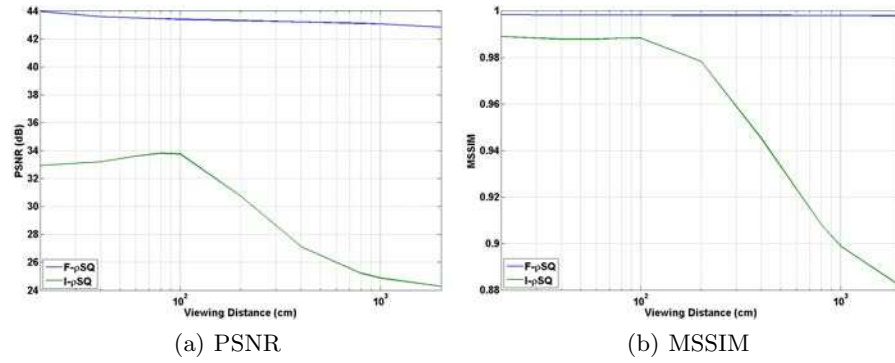


Figure 9. PSNR and MSSIM assessments of compression of Color Images of the CMU image database. Green functions denoted as F- ρ SQ are the quality metrics of forward perceptual quantized images after applying $\alpha(\nu, r)$, while blue functions denoted as I- ρ SQ are the quality metrics of recovered images after applying $\hat{\alpha}(\nu, r)$.

- de Lausanne and Ericsson. [Online]. Available: <http://jj2000.epfl.ch/>
4. T. Mitsa and K. Varkur, "Evaluation of contrast sensitivity functions for formulation of quality measures incorporated in halftoning algorithms," *IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 5, pp. 301–304, 1993.