Gilbert, J., Meynial, X., & Kergomard, J. (1990). The influence of the reed and the reed table on the fundamental frequency of a single-reed system. In *Journal de Physique II*, number C2 in Supplément au Journal de physique, pp. 833 836.

Hirschberg, A., Dane, H. J., Houtsma, A. J. M., Kruijswijk, S. G., van de Laar, R. W. A., Marrou-Maurières, J. P., & Wijnands, P. J. (1990). A quasi-stationary model of air flow in the reed channel of single-reed woodwind instruments. *Acustica*, 70(2), 146 154.

Hirschman, S. E. (1991). *Digital Waveguide Modeling and Simulation of Reed Woodwind Instruments*. Ph.D. thesis, Department of Electrical Engineering, Stanford University.

ICMC (1991). *Proceedings of the 1991 International Computer Music Conference*, Montreal, Canada.

McIntyre, M. E., Schumacher, R. T., & Woodhouse, J. (1983). On the oscillations of musical instruments. *The Journal of the Acoustical Society of America*, 74(5), 1325 1345.

Smith, J. O. (1986). Efficient simulation of the reed-bore and bow-string mechanisms. In *Proceedings of the 1986 International Computer Music Conference*, pp. 275 280, The Hague, Netherlands.

MUSICAL SCORE RECOGNITION OF "DON CUCO EL GUAPO" PIANIST ROBOT

A. Pedroza, H. Vargas, D. Vera and R. Fournier

Dept. of Microelectronics , Universidad Autónoma de Puebla

Boulevard Valsequillo y 14 sur A.P. J–32 Sn Manuel–72570

Puebla, Pue. México

Abstract

Don Cuco el Guapo is the first Mexican pianist robot, which was designed and built at the Departament of Microelectronics of the UAP. The project was based on multidisciplinary participation, where physicists, electronic engineers, computer scientists, musicians and designers converged.

The musical score recognition sytem was implemented through the following steps: frame grabbing, image processing, pattern recognition and interpretation or analysis of scene. The vision system of Don Cuco el Guapo is capable of reading musical score from a template.

Frame Grabbing

Frame grabbing is the process through which a visual image is taken from the three dimensional world. The frame grabbing involves different methods in order to reduce the graphic complexity, increasing the necessary information for object detection and extraction. These methods consist in the precise definition of the object to be captured, that is, what form characteristics does our object have so that the camera set up (focal distance, iris opening and focus) establishes a correspondence between the object (real image) and the plane image (digital image).

An ELECTRIM EDC–1000 camera was used for frame grabbing; it's main characteristics include:

- CCD sensor
- High sensitivity
- Distorsionless image
- Fast response
- Resolution 192(h)x165(v)
- Monochromatic 8 bits
- Spectral range 400–1000 nm

Focal length was taken at one meter, with a variable iris for different illmination conditions. The visual information is converted to electric signal by the sensor CCD. When these signals are sampled and quantized, we obtain a digital image.

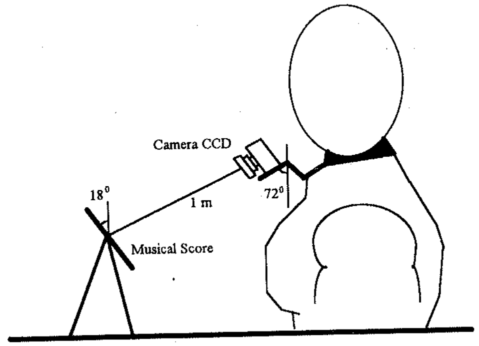


Fig 1: Focal length

The digital image can be represented by the following matrix:

$$f(x, y) = \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0,N-1) \\ f(1,0) & f(1,1) & \dots & f(1,N-1) \\ \dots & \dots & \dots & \dots \\ f(M,0) & f(M,1) & \dots & f(M,N-1) \end{bmatrix}$$

where x and y are discreet variables. Each element of the matrix is called a pixel. For our case the dimension of the matrix is M=165 and N=192.

(a) Musical score format

Don Cuco el Guapo can reach two octaves as seen on the keyboard shown below, along with their representation on the pentagram. From this figure it is possible to observe that a notes position on the pentagram determines a corresponding key, and also the note determines the duration for a given beat.

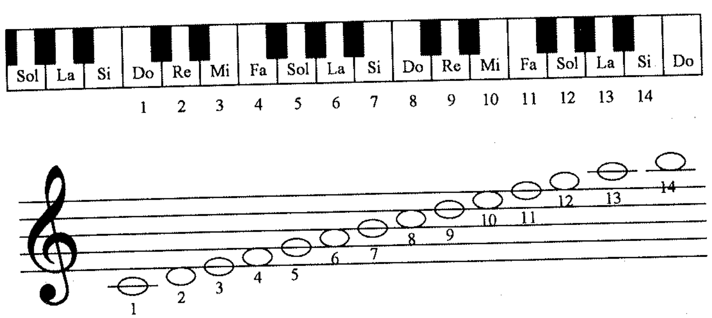
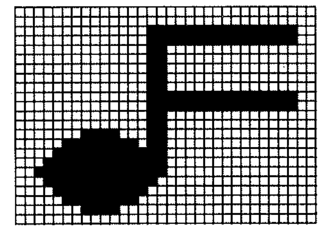


Fig 2: Relation between the keyboard and the working space for each pianist robot arm.

In order to define a musical format that could be adapted to the cameras field of view, the following points were considered:

- Focal length was set equal to 1 meter.
- A field of view equal to 16.4 cm.
- A digital image with 165x192 pixels.
- Each column of the musical score has 14 entries.

All symbols will be defined on the following grid:



23 x 32 (rows x column)

this arrangement provides sufficient definition for later recognition; with consideration having been made of information loss due to image capture and/or processing.

With the chosen parameters, we'll have 6 symbols for each row and 7 symbols for each column, distribuited thruought the field of view. The following figure shows three columns of a beat, and their relation to keyboard positions.

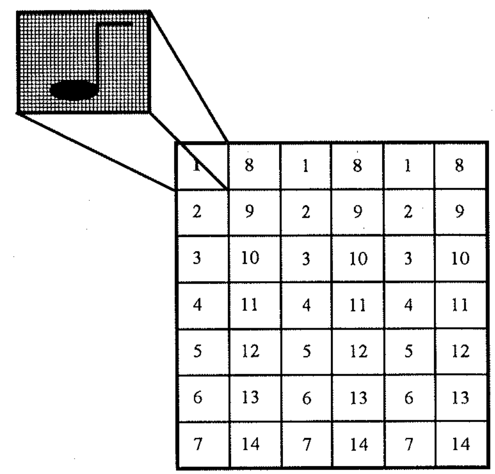
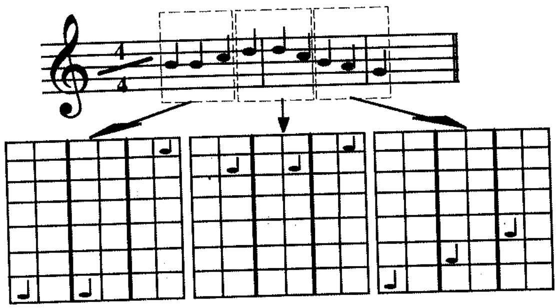


Fig 3: The lines that divide the grid are imaginary. The frame that encloses the field of view acts as a reference for camera perception.

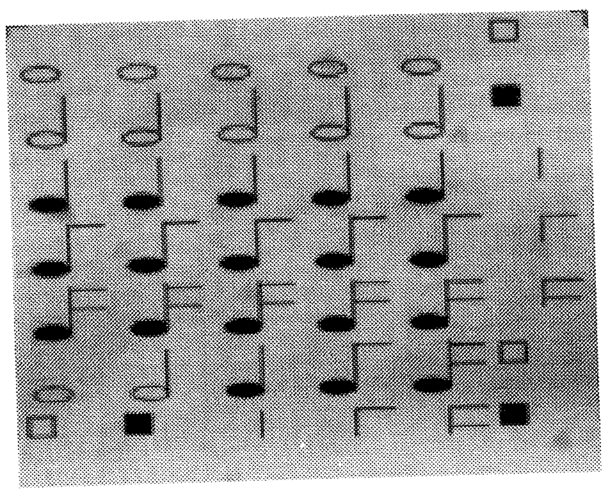
The following is an example of how the process is implemented.

Standard musical score



Musical score adapted to Don Cuco el Guapo's field of view

(b) Example of an image for frame grabbing



This image shows all the symbols allowed by the system, along with their keyboard positions. This image will be the same one used as input to all image processing algorithms.

Image Processing

Image processing is concerned with images generated from existing ones. The new image is the result of applying operations to reduce noise and other artifacts, that may be present as a result of sampling or perturbations in the system.

After analyzing different image processing methods, the Laplace function was chosen for noise reduction. The Laplacian is a second order operator defined as follows

$$L[\,f(x, y)\,] = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial x^2}$$

For digital images the Laplacian is defined as

$$L[\,f(x, y)\,] = [\,f(x + 1, y) + f(x - 1, y) + f(x, y + 1) + f(x, y - 1)\,] - 4\,f(x, y)$$

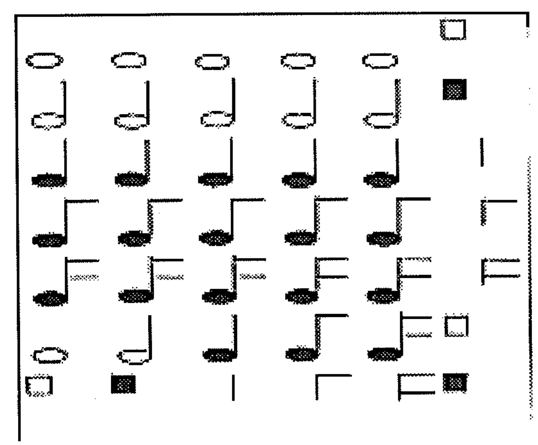
This digital Laplace formula yields zero in zones of constant intensity and in edge ramps, determining if a pixel is on the dark or illuminted side of the edge. Consequently, the Laplacian is used for intensity transitions, and rarely for edge detection.

Equation evaluation may be made using the array shown in fig. 4

0	1	0
1	-4	1
0	1	0

Fig 4: Array used to find the Laplace operator.

The following image results from applying the Laplacian to the input image, being the algorithm that resulted in maximum noise reduction.



Laplace: neighboring surroundings of 3x3.

For our case, a grayscale image is converted to a binary one, by using a non-negative threshold T, where this variable will separate the values in the equation, as follows

$$g(x,y) = \begin{cases} 1 & \text{si} \quad L[f(x,y)] > T \\ 0 & \text{si} \quad L[f(x,y)] \le T \end{cases}$$

Thus, this equation may be seen as a procedure that extracts only those pixels characterized by significant intensity transitions (as set by T). This new binary image will be the entry data for the recognition algorithm.

Pattern Recognition

By Pattern Recognition we understand a process by which it's possible to determine the importance of each feature described in a certain object or phenomenon with respect to it's own characterization, and also to a given class.

(a) Segmentation

The image dimensions are 165x192 pixels (rows x columns), and we divide the image in 6 columns by 7 rows, resulting in a submatrix of 23x32 pixels (rows x columns), leaving 4 rows unnoccupied.

The submatrix is well defined, and it´s area is given by

$$\sum_{i=0}^{n}\sum_{j=0}^{m} A_{ij} \quad \text{where } m = 32 \text{ y } n = 23$$

Evaluation of this submatrix yields the existence or absence of an object in the field of vision of the input grid.

The criterion for determining the existence or absence of an object is given by

$$f(A_{ij}) = \begin{cases} 1 & \text{if} \quad \delta \le \sum_{i=0}^{n}\sum_{j=0}^{m} A_{ij} \le \varepsilon \\ 0 & \text{if} \quad \delta > \sum_{i=0}^{n}\sum_{j=0}^{m} A_{ij} \\ -1 & \text{if} \quad \sum_{i=0}^{n}\sum_{j=0}^{m} A_{ij} > \varepsilon \end{cases}$$

where n=23 and m=32. δ and ε are presence estimators, where d is least and e is most. If the decision function results equal to one, this means a symbol exists as input for that area in the field of vision; if it's equal to 0, this means there is no input, and when it's equal to -1 the input is undefined for the system symbols.

The decision function can determine which note on the pentagram is represented by the object from it's position on the field of view (or on the corresponding grid); thus, if a symbol exists, we determine the note which is it's first parameter (key number). Note duration, which is it's second parameter, will be obtained by the method of description and recognition.

(b) Description

Once segmented, and considering that a symbol was found on an image grid square, then the next step is to make a symbol description, that is, to obtain all the characteristics that define it. Having determined these parameters, we procede with the recognition.

A solution is obtained by dividing the matrix from the corresponding grid, which contains the object, into submatrices of different dimensions (as shown in fig. 5), and which will represent the characteristic features of the object.

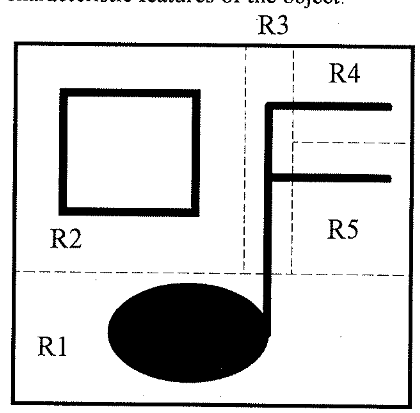


Fig 5: The grid is divided in different submatrices, each one representing an object feature or characteristic. This diagram shows two musical symbols that give an idea of how the different features are discriminated.

All the features for the description procedure have been defined, and they are given below:

- R1 represents the elipse
- R2 represents the round silence or white
- R3 represents the escrow of the symbol
- R4 represents a quaver of the symbol
- R5 represents a quaver of the symbol
- R6 indicates if the elipse is white or black

The result of the description procedure is a vector, which will contain as it's inputs the features of the object, this is

$$\mathbf{v} = (R1, R2, R3, R4, R5, R6)$$

this vector is generated by the following boolean function

$$f(D_i) = R_i = \begin{cases} 1 & \text{if} \quad \delta_i \leq \sum_{i=0}^{n}\sum_{j=0}^{m} D_{ij} \leq \varepsilon_i \\ 0 & \text{other case} \end{cases}$$

where D_i coresponds to the submatrix of the ith feature, i = 0, 1, 2, 3, 4, 5. The submatrix dimensions are given below

For D1, n = 32 and m = 10   For D2, n = 12 and m = 13
For D3, n = 8 and m = 13   For D4, n = 12 and m = 7
For D5, n = 12 and m = 7   For D6, verifies if its white or black

ε_i is a tolerance constant. If the submatrix of the evaluated feature in the function results equal to 1, this means that the object has that feature, if the opposite is true and the result is equal to 0, this means that the object does not have the feature. Then, the description vector will be given by

$$\mathbf{v} = (f(D_0), f(D_1), f(D_2), f(D_3), f(D_4), f(D_5))$$

this vector only yields information for the features found in the object. The decision if the vector has the characteristics of a symbol defined for the system, is given by the recognition algorithm. Consequently, this vector will be the input data for the recognition algorithm.

(c) Recognition

The vector obtained in the preceding section is evaluated in table 1; and will decide if the vector information corresponds to a system defined symbol, ie

1 2 3 4 5 6 7 8 9 10

Fig 5: Sytem defined symbol

Table 1: Object properties.

TOP	R1	R2	R3	R4	R5	R6	VALOR
1	1	0	0	0	0	0	32
2	1	0	1	0	0	0	40
3	1	0	1	0	0	1	41
4	1	0	1	1	0	1	45
5	1	0	1	1	1	1	47
6	0	1	0	0	0	0	16
7	0	1	0	0	0	1	17
8	0	0	1	0	0	0	8
9	0	0	1	1	0	0	12
10	0	0	1	1	1	0	14

It can be clearly observed that all objects have different inputs for their features. The VALUE column represents in decimal code, the value of the object row, and this value will determine if an object belongs to this class or not.

Solution:

Let $\mathbf{v}$ = (R1, R2, R3, R4, R5, R6) the vector that contains all the features of the object. Let f be a function that converts the vector in decimal representation, defined as

$$f(\mathbf{v}) = R_1R_2R_3R_4R_5R_{6\,\mathbf{b}} = n$$

where

$$n \in S \subset N$$

and

$$R_i = \{ 0 \text{ or } 1 \} \quad i = 1, 2, 3, 4, 5, 6.$$

Now, to decide if the object belongs to the class of the musical symbols, use can be made of the decision function

$$g(n) \rightarrow VALUE_i$$

where VALUEi is a value from table 1, this means that if the number n evaluated from function g is equal to one of the values of the table, the object is in the class, otherwise the object does not belong to the class.

Consequently, from this last function we determine the second symbol parameter: note duration. At this point, we recognize the symbol with all it's characteristics.

Also, at this stage, the mechanical limits of the pianist robot are verified; that is, if a chord goes beyond an interval of five adjacent notes, an error is generated, allowing the user to make any necessary corrections.

As the musical symbols are recognized, a doubly linked structure is generated, which will contain the information corresponding to the note or chord, and will be stored in a file.

Interpretation

The performance of an artificial vision system is determined by its capacity to extract meaningfull information from a scene with a wide margin of conditions. Thus, the interpretation, includes all those methods that are related to scene comprehension. In this way, interpretation associates a meaning or an action to a set of recognized objects.

In the pianist robot's vision system, the interpretation algorithm will understand all the musical file, up to the the start of robot's control system, for the the code execution. In the interpretation process , note digitization is made first, that is, notes are assigned to the corresponding fingers and optimization is made of the robot's arm movements. This new code is then translated to another one which will be delivered to the interface.

After this the whole system is activated, and the robot produces computer music, by playing the keyboard with a mechanical sensitivity. In the robot's presentations, science, art and technology converge.

Example of musical score for "Don Cuco el Guapo" pianist robot

Recordando a Bach

Alejandro Pedroza Maléndez

Bibliography

Image Forming Optics - Instrumentals Optics
by Simclair, Douglas.
Image Processing and Computer - aided design in electronics optics
P.W. Hawkes. Academic Press 1973
Image Analysis and Mathematical Morphology
Serra J. Academic Press (London)
Introductory computer Vision and Image Processing
Adrian Low
Otsu, N. (1979) A threshold selection method from grey-level histogram,
IEEE Transaction Systems, Man and Cybernetics, 9(1), 62-66 January.
Advances in Digital Image Processing
Theory, Application, Implementation
Edited by Peter Stucki
Machine - Aided
Image Analysis 1978
Institute of Physics Conference Series Number 44
Pattern Recognition Introduction and Fundations
by Sklansky Jack
Pattern Recognition Principles
by Tou, J.T. & Gonzáles, R.C.
Pattern Classification and Scena Analysis
Duda, Richard O. & Hart, Peter E., John Wiley & Sons
Inc. 1973
Modelos Matemáticos para el Reconocimiento de Patrones
José Ruiz Shulcloper
Instituto de Cibernética, Matemática y Física
Academia de Ciencias de Cuba
Robot Vision
The MIT Electrical Engineering and Computer Science Series
Berthold Klaus, Paul Horn
Fundamentals for Control of Robotics Manipulators
Antti J. Koivo
Robotica: Control, Detección, Visión e Inteligencia
K.S. Fu, R.C. González, C.S.G. Lee
A. Bazerghi, A. A. Goldenberg, J. Apkarian
"An Exact Kinematic Model of PUMA 600 Manipulator"
IEEE Transactions on Systems, Man, and Cybernetics,
Vol. SM-14, No. 3, pp 483 - 487, May - June 1984
C.P. Neuman and J.J. Murray
"Symbolically Efficient Formulations for Computational Robot Dynamics"
Journal of Robotic Systems, Vol. 4, No. 6, pp 743 - 769 1987.
M.W. Walker and D.E. Orin

"Efficient Dynamic Computer Simulations of Robotic Mechanisms"
Journal of Dynamical System, Measurments and Control, Vol. 104,
pp 205 - 211, 1982.
DC Motors, Speed Control Servo Systems
An Engineering Handbook
prepared by Electro - Craft Corporation, USA

MIDISCAN

the program for reading and processing musical notation

WLADYSLAW HOMENDA(*)(**), ALFREDO CRISTOBAL(**), MERCEDES ARELLANO(**)

Department of Electronics and Telecommunication, Applied Physics Division
Centro de Investigacion Cientifica y de Education Superior de Ensenada
Km. 107 Carretera Tijuana/Ensenada, Ensenada, B.C., Mexico 22860

ABSTRACT

The paper presents problems related to automated recognition of printed music notation. Music notation recognition is a challenging problem in both fields: pattern recognition and knowledge representation. Music notation symbols, though well characterized by their features, are arranged in elaborated way in real music notation, which makes recognition task very difficult and still open for new ideas. On the other hand, the aim of the system, i.e. application of acquired printed music into further processing requires special representation of music data. Due to complexity of music nature and music notation, music representation is one of the key issue in music notation recognition and music processing. The problems of pattern recognition and knowledge representation in context of music processing are discussed in this paper. MIDISCAN, the computer system for music notation recognition and music processing, is presented.

Keywords: music notation recognition, knowledge representation, music representation, MIDI format.

1. INTRODUCTION

There are many ways in which computers have been involved in the world of music. One is the score edition, where musicians can develop a score using an editor to produce a digital output file. This strategy might be a good idea, but it is a non-traditional way to develop a score. Another computer application is music processing that can help musicians in music creation process: automatic composition of music, analysis of musical style and so on.

(*) ***On leave from:*** Institute of Mathemaics, Warsaw University of Technology, Warsaw, Poland

(**) ***Correspondence to:*** Wladyslaw Homenda, Mercedes Arellano, Alfredo Cristobal,
- CICESE Research Center / Applied Physics, P.O.Box 434944 San Diego, CA 92143-4944, USA
  e-mails to: {homenda,arellano,asalas}@cicese.mx
- Wladyslaw Homenda, Instytut Matematyki, Politechnika Warszawska, pl. Politechniki 1, 00-661 Warszawa, Poland
  e-mails to: homenda@plwatu21.bitnet