

INTEGRATED SYSTEM FOR MECHANIC BAND CONVEYORS WITH HYBRID MOTOR, TV CAMERA AND ROBOT

Radu PENTIUC, Ph.D.,Eng., Stefan-Gheorghe PENTIUC, Ph.D.,Eng.,

Adrian GRAUR Ph.D.,Eng., Mariana MILICI Lecturer Eng.

University "Stefan cel Mare" of Suceava, Romania
Faculty of Electrical Engineering

radup@eed.usv.ro, pentiuc@eed.usv.ro adrian@eed.usv.ro dam@eed.usv.ro

Abstract: The aim of the paper is to present a system used for the unloading of the mechanic band with low speed and carrying heavy materials. The movement is assured by an hybrid electrical induction motor able to develop a sufficient force according to the speed of the band and the mass of the transported blocks. The blocks are uncharged and pushed into a stack by a robot, in fact a manipulator with 6 liberty degrees controlled by a computer. Two TV cameras are connected to the computer. Based on the image analysis the computers synthesises the commands for the robot.

Keywords: hybrid electrical induction motor, digital image processing, robot, automation

INTRODUCTION

The mechanic conveyor bands are widely used in constructions and in the industry of construction materials.

Constructions band conveyors can work in the environment, in a tuff climate activating against the electric and mechanic elements. The sand, the dust, the frost and the humidity have a negative influence on the equipment's working and affects especially the function of the reduction gears of revolution.

The main inconvenient of the low speed mechanic band conveyors, trained up with electric

motors, usually asynchronous, is that the mechanic reduction gears of revolution, coupled with the activating engines, present a high incidence of the defects by reducing the period of working of the technological equipment because of the frequent interventions for reparations.

This paper describes some main aspects of an integrated system for unloading a conveyor band (see figure 3), presenting the control of this operation by a robot guided by computers processing the images captured by two cameras and scanning the outputs of an automate controlling the hybrid electrical induction motors that trains up the conveyor band.

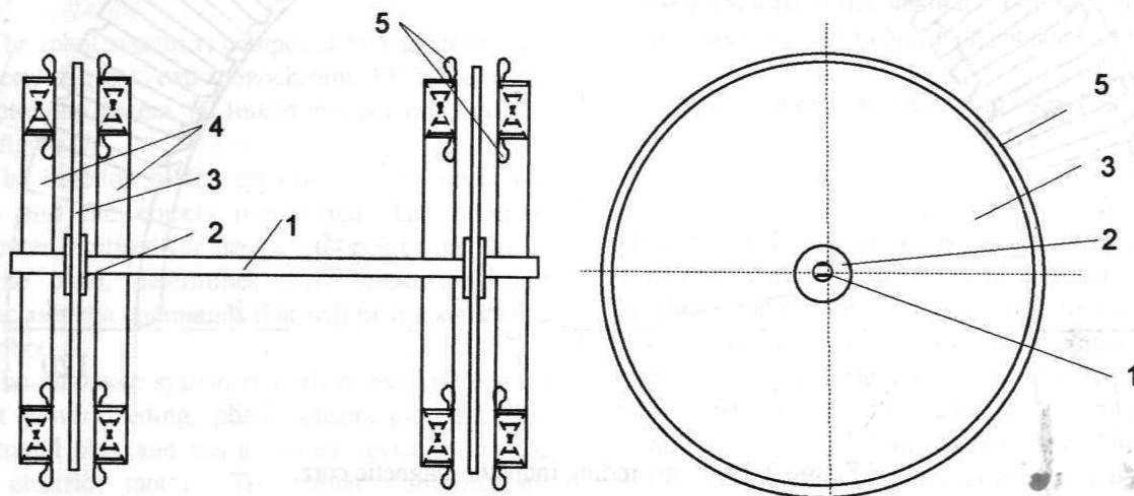


Figure1. Activating system at low speed with EIHM for mechanic band conveyors.

1. ELECTRIC ACTIVATING SYSTEM AT LOW SPEED WITH HYBRID MOTOR FOR MECHANIC BAND CONVEYORS

It is introduced a new modern activating system at low speed, equipped with an electric induction hybrid asynchronous motor (EIHM) [3] for the mechanic band conveyors used in constructions.

1.1. Particularities of an activating system with EIHM

The technical problem that is resolved by our installation consists of the expulsion of the reduction gear of revolution from the action system by directly training the band conveyor with low speed EIHM, this way resulting the reducing of the off-times for the repairing of the defects of the reduction gears, reducing the costs of the conveyor equipment, the weight of the equipment, diminishing extra expenses for maintaining, when the movement from the motor to the band is done directly, without using a reduction gear, the power on the installation goes higher, the transporting installation can be also built in a modulate way with doubled groups of EIHM, at every drum made in order to activate the band. The conveying band trained directly by the EIHM, in the new proposed variant moves away the up mentioned

disadvantages because the activating system is realised with EIHM, which develop a low speed revolution motion [5]; this fact permits the exclusion of the cinematic chain of the mechanic reduction gear of revolution.

An EIHM is recommended to be built with $p \geq 4$ poles number. After the radius R dependent, pole pitch τ , the tangential synchronous velocity v_i is computed using the primary supply tension frequency f_i :

$$v_{i(R)} = 2f_i \tau_{(R)} \tag{1}$$

where $v_{i(R)}$ and $\tau_{(R)}$ are values dependent segment circle inductor radius.

In the base of the maximum radius of inductor R_{max} and the angular speed ω , the tangential synchronous velocity $v_{i(R_{max})}$ is calculated.

1.2. Compounds parts of an EIHM with bilateral toroidal inductors

The EIHM has its primary winding on the toroidal inductors placed bilateral from the disc - made of conductive material - together with the activating drummer of the conveying band, this disk is also the inducted with the secondary winding of the hybrid motor..

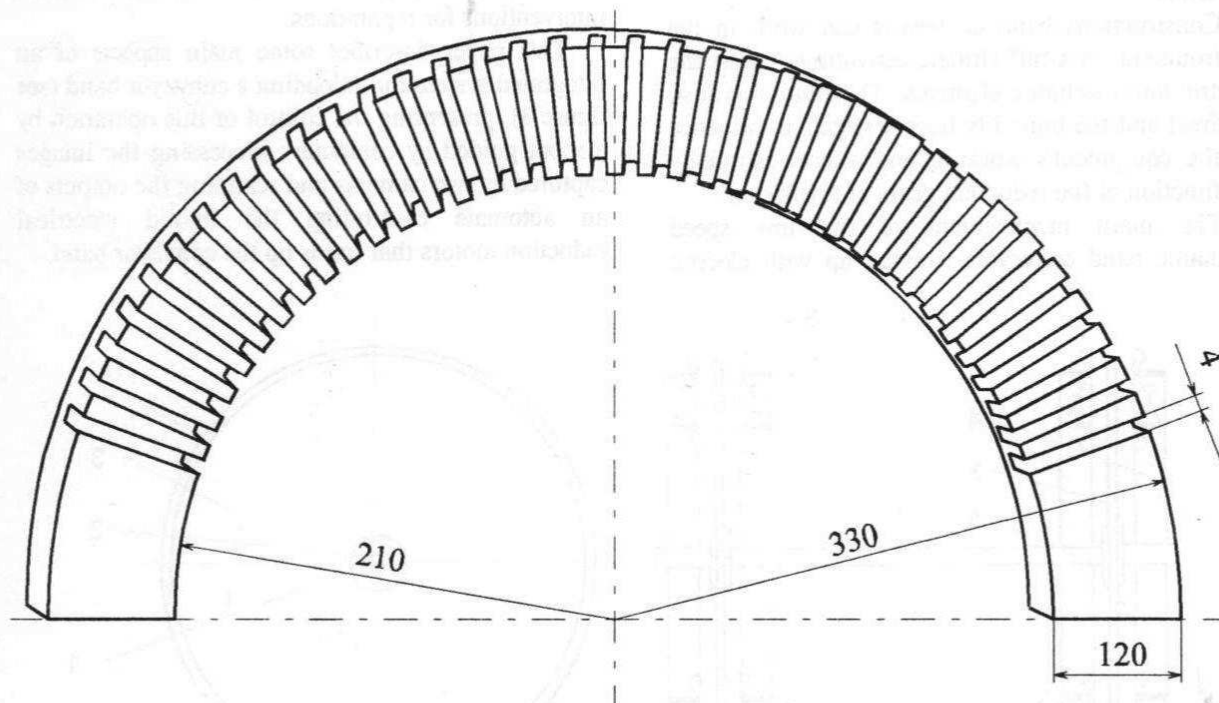


Figure 2. Built unwinding inductor magnetic core.

In fig.1 is presented an activating system with EIHM for conveying bands used in constructions, mining exploitations, in quarries, in the metallurgical

enterprises, in port fitting outs, for cereals and food business, for rolling stairs or other systems of transport a 1 drummer on which the rolling band of

the conveyer is held, by using the elastic couple 2, there are linked on the lateral sides two disks made of electric secondary winding of the hybrid motors. On both sides of every disk 3, there are two fixed toroidal inductors 4, on which is placed the distributed primary winding 5, or in the form of a ring.

In order to improve the performances of this electric machine, the disk of conductive material can be played of the middle, a sandwich-like structure, by using super magnetic material (silicosis sheet iron or other iron -magnetically material).

The toroidal inductor of the EIHM may have an opening segment as a semi-toroidal (semicircle), the construction of the primary magnetic core without winding being represented by fig.2, or as a complete semicircle opened at 360° , case in which the engine becomes an asynchronous one with a rolling disc.

The core in fig 2 has been projected, built and tested by the author in lab by using a testing.

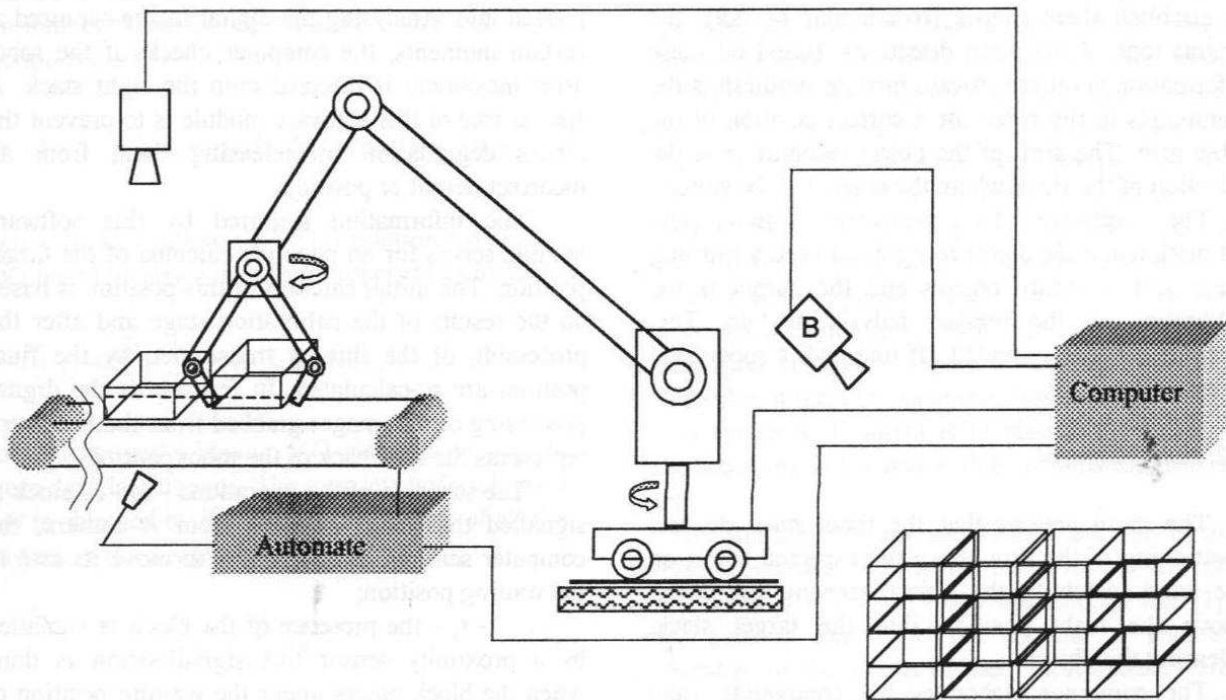


Figure 3. Integrated system for conveyor band unloading.

2. ROBOT CONTROL SYSTEM

The robot system is composed by one robot with 6 liberty degrees, two monochrome TV camera and an automate system, all linked to a computer system (see figure 3).

The function of the camera A is to survey the band and the objects transported. The program computer digitised the image and locates the blocks on the band, determines their orientations and synthesises the commands that will be transmitted to the robot.

The automate system centralises the information about power feeding, photo sensors placed in the horizontal plan and has a special device to control the electric motor. The binary information, especially from photo-sensor are redirected to the computer.

The B camera surveys the stack made by the robots. The monochrome image is grabbed and

digitised by the computers. The results of the image analysis are used in the elaboration of the commands for the robot in order to build a right stack of blocks.

2.1. Acquisition and analysis of the band surface images

The surface of the band is surveyed by the monochrome TV camera named A in the figure 3. It is assumed that exists a sufficient contrast between the blocks and the surface of the band. In most cases this contrast is assured by the black colour of the band and the grey colour of the blocks (BCA). The area surveyed by the camera is artificially illuminated with constant parameters. For these reasons, the use of a monochrome camera is perfectly justified, but this is not an imperative requirement. The main stages of the image processing are: digital image acquisition, binarisation, object labelling, filtering, object

analysis (shape, orientation, geometrical parameters estimation).

In order to have a certain robustness face to the variations of band colours or blocks, illumination conditions, the binarisation is made with an algorithm based on pattern recognition technique [2]. The binarysed image permits the execution of a quickly procedure for labelling objects and detecting a number of ROIs (Region of interest) containing a single object. These ROI's may be analysed one by one, leading to a high speed of processing. Some ROI that are not eliminated in the filtering process contain relevant object which are analysed in order to establish their shapes (rectangular or not), the orientations of the main directions. Based on these information another software module synthesises the commands to the robot for a correct position of the robot arm. The aim of the object recognition is the selection of the stack where the object will be stored.

The software for geometrical parameters estimation and the object recognition need a learning stage with standard objects and the output is the calibration of the image analysis module. This learning stage is executed off-line and is supervised by an operator.

2.2. Robot Control

The main actions that the robot must do are: positioning of the arm above the expected object on the band, catching the object, moving the object above the right position into the target stack, releasing the object.

The computer elaborates the commands and transmits them by a serial link to the robot. The fields of a robot command must contain polar co-ordinates in the robot's own co-ordinating system. In the frame of the whole control system of the robot, including TV camera, there may be distinguished 7 co-ordinates systems:

- object surveying system (S1) - measuring units: cm
- graphic system for the A TV camera (SA) - measuring units: pixels
- object catching area system (S2) - measuring units: cm
- robot co-ordinates system - polar co-ordinate for each arm (S3), measuring units CAN
- stacks systems (S4) and (S5) - measuring units: cm (one for the stack with standard objects and one for the non-standard ones)
- graphic system for the B TV camera (SB) - measuring units: pixels

These systems assure the flexibility of the operation. The co-ordinates transformation relations is based on coefficients computed during the learning stages (calibration). For example, in order to have a correct transformation between the co-

ordinates in S2 and S3, the arm of the robot is moved, by a manual command of the operator, in the origin of the S2 and in a point placed on the Ox axis of the S2. As answer to a special command emitted by the computer the robot transmits the co-ordinates (in S3) of its arms.

2.3. Stack building

The arrangement of the blocks into the stacks is supervised by the computers. The B TV camera surveys the stack where the standard objects are pushed into. Analysing the digital image captured at certain moments, the computer checks if the robot arms movement is directed onto the right stack. A special role of this software module is to prevent the blocks degradation by releasing them from an incorrect height or position.

The information acquired by this software module serves for an adaptive calculus of the target position. The initial calculus of this position is based on the results of the calibration stage and after the processing of the images transmitted by the final position are re-calculated. In few words the digital processing of the images grabbed from the B camera represents the feed-back of the robot control.

The schedule of the operations - to - a block is signalled the image analyses from A camera; the computer send an order to robot to move its arm in the waiting position;

- t_1 - the presence of the block is signalled by a proximity sensor this signallisation is done when the block enters under the waiting position of the robot arm; the commands to catch the block are transmitted to the robot

- t_2 - the robot begins to the block to the target stack.

- t_3 - another series of commands based on the image analysed the B camera is launched to the robot

- t_4 - the robot release the block on the stack.

At this final point a new block enters in the area surveyed by the A camera and a new cycle begins. The computer control the speed of the band by estimating the frequency of the objects on the band.

3. IMAGE PROCESSING AND RECOGNITION

3.1. Image processing related aspects

During the acquisition of monochrome images different noise sources may diminish the quality of digitised image. It may be distinguished two main types of noise sources: deterministic (e.g. optical system imperfections) and stochastic. The restoration

of a monochrome digitised image is based on the following model of the image degradation and image restoration processes.

The original image is represented by the function $o(x, y)$ giving the brightness value in the point of coordinates x and y . We assume that the degradation was done by a filter with the transfer function $H(u, v)$ and an additive noise $z(x, y)$. For the restoration is responsible a filter with the transfer function $R(u, v)$. The restored image function, $f(x, y)$, must approximate the function $o(x, y)$ representing the original image. The degraded image function $d(x, y)$ is assumed that is a sum of the result of a space invariant convolution operator $h(\cdot)$ and the noise function $z(x, y)$:

$$d(x, y) = \int_{-c}^c \int_{-c}^c h(x, y; \alpha, \beta) o(\alpha, \beta) d\alpha d\beta + z(x, y) \quad (1)$$

where c is the window size. This relation may be re-wrote based on space invariant properties as follows

$$d(i, j) = \sum_{k=-c}^c \sum_{l=-c}^c o(i-k, j-l) h(k, l) + z(i, j) \quad (2)$$

A neural networks having $N.G$ neurones is used to restore dgradeate images. The grey level of each pixel $r(i, j)$ is obtained by the sum of the outputs of all the neurones in the assigned group.

$$r(i, j) = \sum_{k=1}^G v_{ij}^k \quad (3)$$

where v_{ij}^k is the output of k -th neurone in the group of G neurones assigned to the pixel (i, j) . This output is obtained by applying a threshold function $g(u)$ to the total input of the neurone. All the neurones are assumed to be interconnected by symmetric links.

We consider a monochrome image digitised in G grey levels and $N = x_{max} \cdot y_{max}$ pixels. After the binarisation follows the phase of objects isolation that is based on a convex compounds labelling method. It must be underlined that every connecting area of pixel is interpreted as an object marked by the same label.

After the identification of the object representing a block it must be determined its orientation. The main axes of the object are computed using an analogy to the samples points in an stochastic process, and were approximate by a regression line

$$y = ax + b \quad (4)$$

where the coefficients a and b are determined by the least square method using the pixels points in the object waiting area image. Their values minimize the function $F(a, b)$, computed with the formula:

$$F(a, b) = \sum_{(x_i, y_i) \in \text{Object}} (y_i - ax_i - b)^2 \quad (5)$$

where the *Object* stands for the set of the pixels in the object image representing the carried block.

Having the coefficients a and b the main orientation of the block is determined.

3.2 Recognition of standard objects

The significant objects identified on the conveyor band may be standard object (whole blocks), pieces of a fragmented block etc. the standard objects recognition is necessary to send to the robot the order that will move the objects in the standard objects stack, and the other objects in the other stack. The recognition process is based on a the concept of parallel separated patterns [4]. A pattern is described by an array of features or a pattern may be viewed as a point in a n -dimensional space. The similar patterns can be grouped into a finite number of sets. We suppose that these sets are disjoint and we name them classes; their representation in the pattern space consists in a number of regions that may be separable or not. For example a class contains

To recognize a pattern is to classify them into a class, according to the region that contains the pattern point.

The similarity measure used in this case may be an inverse dependence from a distance measure between two patterns $x, y \in E$. In many applications like [1], arises patterns that have not all the n features measurable, or the results of their measurements are strongly affected by noise. Suppose there are only k out of n features measurable for all the patterns. Then we will work in a projection of pattern space. More difficult is the situation that the patterns have not the same features incomplete measured. In this situation the usual solutions can be:

- a) isolating incomplete patterns by:
 - elimination from the set E ;
 - classification of the incomplete patterns in a special class [1].
 - b) restoring the missing values by:
 - substitution of the missing features by the means of the known values of the features of the other ones patterns;
 - computation of missing features by the aid of a regression function determined from complete patterns.
- A major disadvantage of these methods is the requirement of a large mount of data to make possible the elimination or restoring the missing data. But in many cases it is not necessarily to compute the missing features. More important is to gain the final objective of the pattern recognition problem. In the supervised learning context the system receives as inputs data a set composed of prototypes with known classification composed of prototypes with known classification:

$$T = \{ (x, C_x) / C_x \subseteq E \text{ and } x \in C_x \} \quad (1)$$

where T is called training set, with C_x a class in the partition P over E . Considering T , the recognition system must assign a class of P to any new pattern y . This assignment may be done directly by a search in T , (pattern matching) or by a classifier that functions as follows:

- for K classes are K functions g_i ;
- it classifies the new pattern y in C_i class iff

$$g_i(y) \geq g_j(y) \quad (\forall) i, j \leq K \quad (2)$$

The functions $g_i(x)$ may be linearly or non-linearly.

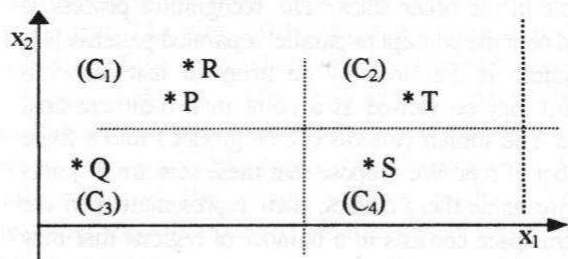


Figure 4. Case of four patterns classes inseparable in an one dimensional feature space (x_1 or x_2).

The quickly classifier is a "third order classifier" type. Clearly, a classifier obtained for complete patterns is inapplicable for patterns with missing features. An expansive solution is to obtain several classifier g_{ij} . For n features it needs a number of N_c classifiers

$$N_c = C_n^1 + C_n^2 + \dots + C_n^n \quad (3)$$

It is important to observe that the separability of the classes may be not maintained in the projection space (figure 4).

Here we had supposed that the pattern space is a bi-dimensional one and we have four classes labelled by C_1, \dots, C_4 . In this case if a feature is missed, it is not possible to separate the pattern in four classes, but only in two (e.g. if x_2 is missing, it may observe a class containing the patterns from $C_1 \cup C_3$ and a class corresponding to $C_2 \cup C_4$). The separability analysis is accomplished during the determination of classifiers, when the cases of uncertain separability are rejected for automate recognition. The corresponding objects may be falses or non-standard objects.

The classifiers determined in the learning phase are algorithms with polynomial time complexity very appropriate to the decision making process running in the time that the object is moved from the A TV camera visual area to the area where it will be picked up by the robot arm.

CONCLUSIONS

The installation presents the following advantages:

- the direct transmission of the movement between bilateral toroidal inductors which represent the main fitting of the action motor and the inducting solidary disk by mechanically coupling with the action motor and the inducting solidary disk by mechanically coupling with the activating drummer of the conveying band;
- the expulsion of the reduction gears of revolution which diminishes the global efficiency of the equipment;
- easy and cheaper maintenance;
- protection against shock and trepidations of the mechanic reduction gears of revolution;

The system assured also the increase of the reliability and liability.

The correctness of the robot movements are dependedent of the calibration stages. These stages must be done only if one component of the system (conveyor, robot, cameras, stacks) changed its position.

REFERENCES

- [1] Celeux G., Diday E. et all *Dclassification automatique des donnees*. Dunod, Paris, 1989 .
- [2] St.Gh.Pentiuc,R.Prodan,R.Haliuc,L.Iurescu. "Monochrome Image Binarization Method Based on a Pattern Recognition Technique". *The II International Conference on Microelectronics and Computer Science*, Vol.II, pp. 203-206. Technique Publishing House, Chişinău Moldavia, 1997.
- [3] R. Pentiuc, L. Cantemir. "Considerații asupra alterării câmpului magnetic datorită efectului de bord la motoarele hibride cu inductoare toroidale." *First International Conference SIELMEC'97*, Vol. I, pp. 302-305. Technique Publishing House, Chişinău Moldavia. 1997.
- [4] St. Gh. Pentiuc. *Aplcatii ale recunoasterii formelor in diagnosticul automat*, Ed. Tehnică, Bucureşti. 1997.
- [5] R. Pentiuc. "Le calcul des reactances de magnetisation des enroulements de tipe anneau pour les moteurs hybrides", *The Annals of the "Ştefan cel Mare" University of Suceava*, Electrical Section, No.8, pp. 50-55. 1998.

Towards a Natural User Interface for Comprehensive Support of Conceptual Shape Design

György Kuczogi

Zoltán Rusák

Imre Horváth

Faculty of Design, Engineering and Production

Delft University of Technology

Jaffalaan 9, 2628 BX, Delft,

The Netherlands

TELEPHONE: +31 15 278 1196

FAX: +31 15 278 1839

E-MAIL: g.kuczogi@io.tudelft.nl

ABSTRACT

The lack of possibility of creative communication in the conceptualization phase of the design process significantly decreases the effectiveness. The aim of this research is to find effective non-traditional shape conceptualization techniques.

The paper presents an analysis of the activities in the conceptualization process and the man-machine communication. It also shows how to evaluate and select the most efficient shape conceptualization techniques.

1. INTRODUCTION

1.1. Problem statement

It is known for long time that the lack of creative communication between humans and CAD systems is one of the most significant problems that are hindering efficiency. An adequate interface should allow designers to concentrate only on their specific design task, without paying attention to how communicate with the computer. There exist solutions supporting specific activities of native human communication, but we cannot find any system on the market yet, which wraps up all issues relating the conceptualization design tasks.

1.2. Hypothesis

In our opinion, the ideal computer aided conceptual design (CACD) system would be similar to the one bounded by the dashed rectangle in Figure 1. The ideal

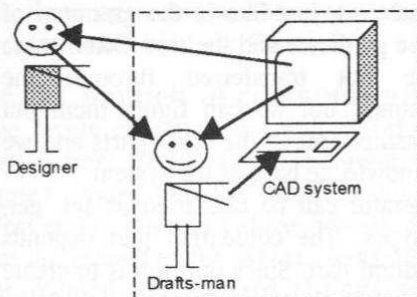


Figure 1 The ideal CACD system should behave like a smart drafts-man and one of the recent CAD software together

system should have an intelligent user interface, which understands the high level natural communication of the human being, corrects the small mistakes without notifying the user, and advises and proposes solutions.

1.3. Aims of the research

Our long-term aim is to create a design system, which supports the complete shape conceptualization process by integrating technologies based on the specific needs of the application field.

The objective of the paper is to (a) analyze the shape conceptualization process, (b) discover the significant elements and influencing factors of the communication in the conceptualization process, and (c) show how can we arrive at more effective shape communication techniques.

1.4. Research methodology

Research into creative communication needs a specific methodology. This paper covers those steps that are indicated in *italic* of our research methodology described below.

- *Definition of application field*
- *Analysis of the activities of computer aided conceptualization*
- *Characterization of the input*
 - *Contents of communication*
 - *Channels of communication*
 - *Level of abstraction in communication*
- *Selection of appropriate techniques for user interface in particular application field*
 - *Elaboration of the aspects of evaluation*
 - *Evaluation/ranking of the categories of techniques*
 - *Selection of the techniques from the top-ranked categories*
- *Composing a framework of an effective system interface for a particular application field by integrating the top-ranked technologies.*
 - *Proposing algorithms for the selected technologies*
- *Implementing the technologies*
- *Evaluation and testing*

2. INVESTIGATION OF THE CONCEPTUALIZATION PROCESS

As a starting point of the analysis of the activities related to the application field, we investigate the conceptualization process in general. Afterwards, we look at the specific activities of shape exteriorization.

2.1. The spiral of the conceptualization process

The conceptualization process is a loop formed by steps for creation/modification and evaluation of concepts. In the case of computer supported conceptualization, we can describe the loop by three partial models:

- the mental model of design concepts,
- the computer model of design concepts, and
- the mental model of the model of the computer (we might call it the designer's view on the computer model).

These three models create a kind of loop based on their contextual dependencies. By introducing one more dimension of investigation for the conceptualization loop, we can represent the evolution of concepts and models as a spiral (Figure 2). The diameter of the spiral shows how close the three models are to each other semantically. During the conceptualization process, each model usually changes and gets closer to the others. When a given computer model resembles the design concept so much that the designer is satisfied with it, the process is finished.

Let us have a closer look at the steps of a general conceptualization process. Consider the following scenario:

1. The three models are given. (The computer model may be an empty model at the beginning.)
2. The designer evaluates the models and the design concept:
 - He compares his view about the computer model to the original mental model to judge how accurately is the concept represented by the computer.
 - Based on his view about the computer model, he investigates the design concept, whether the concept really fulfills the requirements.
3. The designer communicates the needed changes to the computer.
4. The computer modifies its digital model.
5. The computer communicates the modified model to the designer. (The loop continues with the step 2.)

2.2. Activities of the conceptualization process

There are three activities in this scenario mentioned above, where the computer has an important role. Although it needs much deeper investigation to see how we can speed up the conceptualization process, in this subchapter we limit ourselves to the essence of the activities.

The first activity is when the designer communicates the required changes of the model to the computer. The

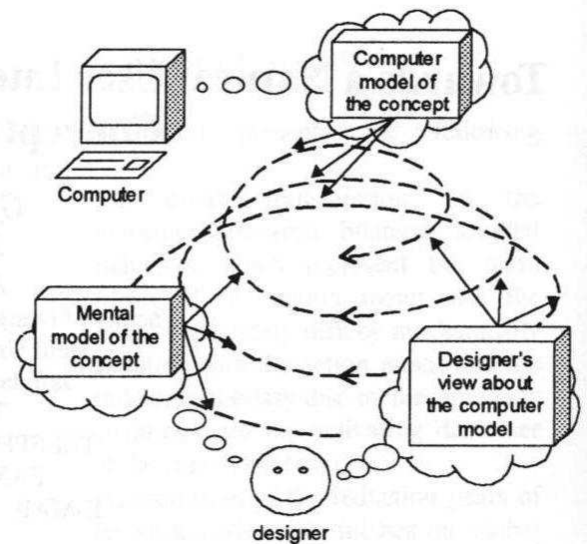


Figure 2 Evolution of models in the conceptualization process

computer must interpret and react on the communicated contents. We will go into a more detailed analysis of it in the next chapter.

The second activity is modification of the computer model. The effectiveness and opportunities of modification are determined mainly by the applied representation. The computer has to be capable to handle the incomplete and vague concepts, which is one of the most dominant characters of conceptualization. We have decided to use discrete model representation. The paper of Rusák et al [1] discusses it in detail.

The third activity is communication of the computer model to the designer. There are several existing output systems on the market. We can find numerous solutions in the field of scientific visualization, for example, displaying the model in virtual reality, producing photo-realistic pictures etc. According to our observations, the output systems tend not to be the bottleneck of the computer support of the conceptualization process.

3. INVESTIGATION OF THE INPUT COMMUNICATION

3.1. Structure of communicated instructions

The communicated instructions have a communication independent structure. Each of them has a procedural part, and a contextual part, just like in the grammar of natural languages: the predicate and the rest. Often some of the parts are not transferred through the communication channels, but we can figure them out based on the relationships among the other parts and we need to rely on the knowledge base of the system.

Each procedural operator can be classified as set, get, create, or destroy types. The contextual part depends largely on the procedural part. Since our aim is to create a shape model, the classification of elements of the contextual part of communication is exactly the same as with the regular geometric models of CAD systems. We

distinguish geometry (morphology and scale), topology, identification information, location, and attributes. Depending on the category of a particular element of a user instruction, it may have a most effective device to communicate it.

3.2. Dependence of the aspects of communication

Referring to the first activity in the earlier mentioned scenario, if the designer is not satisfied with the digital shape model, he wants to change it by communicating one or more instructions to the computer. He can do it in several ways. He may choose from the available modeling operators, communicate his choice on different abstraction levels, and select different kinds of communication channels. We can conclude that the three aspects of communication, i.e. (a) content of communication, (b) the channel of the communication, and (c) the level of abstraction, are independent from each other.

3.3. Contents of communication

In order to describe a domain of the desired shape or modify its attributes, location etc., we use operators. The operators are determined by their types (procedural part) and parameters (contextual part). We have operators to create, modify, destroy, or get any of the parts of the geometric model. The most influencing operators are the shape modification operators, therefore we present the categories of them based on the theory they apply.

- The create operator is very simple. It does not determine the boundary of the object on a procedural way, but with a large set of discrete points. E.g. this can be applied by using data gloves or 3D scanner.
- The copy operator makes a copy of the boundary or a part of the boundary of a particular object. It facilitates not only the regular copy/paste type of operations, but to apply metaphors, as well.
- We have a number of volumetric operators, which use the set theory, e.g., the "mix" operator [3] uses the union. They are used by all cutting type of modifications.
- Weight operators can modify the shape like parameterized features by weight functions.
- Physical based operators [3] use the theory of finite element methods. We can modify our shapes with these operators as they were physical objects, we can bend, twist etc. them.

3.4. Channels of communication

The levels of communication of information are well known from the literature. The number of levels is still argued issue, but the main points are established. Weaver [2] discusses three aspects of communication that are related to the syntax, semantics, and effects. In our opinion, the selected channel and coding of the communication concerns the physical, statistical, and syntactical levels of communication. The channel can be

voice, keyboard, mouse, hand gesture etc., which form the physical layer of communication. We can apply different coding, e.g., in the case of voice based information transfer, we can use natural languages. The coding covers the statistical and the syntactical levels. What we get at the end of the processing of this input is supposed to be syntactically correct information. No matter what sort of channel or coding we choose, the same semantic elements (the type and parameters of the modeling operators) have to be transferred through. Therefore, it has to be possible to convert the input to a channel independent syntax. We can express exactly the same by hand gestures through a sign language as by talking, although both the channel and the coding are different.

3.5. Level of abstraction

We can express the same intent on the same channel several different ways. The basic difference appears in the level of abstraction. This has also something to do with the levels of information communication.

For instance, we can use the lowest syntactic level. In this case, the syntactic description that we get from the processing the input signals is executable without interpretation. This means that the user must follow a prescribed syntax, and each syntactically complete instruction can directly be connected to a modeling operator.

A more advanced level is the semantic level, where we can attach semantic meanings to the content. In our case, the semantics refer to the modeling operators. This level of the communication also allows the user to apply casual syntax, for instance, he may even skip the required parameters. However, the semantic meaning of the communication must be unambiguous. The reason of it is that the computer has to figure out, which modeling operator was intended to be used. Applying the semantic level of communication may make the use of the shape conceptualization system more understandable for the user.

The third level of abstraction is the pragmatic level, when the user does not need to specify at all, the appropriate modeling operator to achieve the desired shape. He can declare requirements for the concept, and the computer has to come up with proposals, how to fulfill. For instance, the user can declare that he wants a more streamlined shape, and the computer is supposed to propose changing the curvatures of specific regions of the shape.

The fourth level is the apobetical level. It relates much more to the human being than to the shape conceptualization. We consider here feelings, motivations, jokes, etc.

4. SELECTION OF INPUT TECHNIQUES

A particular combination of the communication channel, level of communication, and a clearly defined set of modeling methods is called here shape conceptualization

Table 1 The needed knowledge categories for a given communication level

Level	Knowledge required
Physical	Nothing
Statistical	Coding rules
Syntactical	Type information of the categories of the elements of the geometric models Type information and algorithm of operators Relationships between the types of operators and categories of elements
Semantic	Predefined objects, including semantic parts of the objects (e.g., front, top), what the user can refer to Operators with predefined conditions, which has pictorial meaning for the user Predefined constant attributes, e.g., "red" instead of numbers
Pragmatic	Relationships between the functional requirements and theories of shape modifications
Apobetical	Human feelings, motivations, intentions

technique. The primary issue is to find the most efficient input techniques for our shape conceptualization system. In order to achieve our goal, we can create a three-dimensional matrix, one aspect of communication represented on each axis. We can indicate several techniques at each cells of the matrix. Obviously the techniques that we put the same cell of the matrix, use the same theory of their operators, the same channel and the same level of communication. (For instance, saying "Take the intersection of the two objects" or "Take the union of the two objects".) There can also be minor differences, e.g., in the expression mode, or some operators may require different number of parameters etc. Although, this may influence the efficiency, we can assume that the efficiency of the techniques that share a cell is nearly the same. Therefore, we can conclude that it is not necessary to name and define all the techniques before evaluating them. It is sufficient to evaluate only the cells of the matrix, and we can define the actual techniques later.

The evaluation procedure is twofold. On the one hand, the theory of the operators and the level of the abstraction determine the actual syntax to be transferred. On the other hand, a particular channel of the communication can be more suitable for transferring the given syntax than other channels. Let us elaborate more about the two issues:

In Chapter 3.3., we have introduced the five theories of shape definition/modification that we want to use. Therefore, we have to choose, which level of abstraction is the most appropriate for the particular theories. Although, we have to send significantly less data to the system by increasing the level of communication, it is not always the case that the higher level we communicate the faster the result is achieved. We have to consider that using high level communication

accompanied by uncertainty. The level of communication is proportional the knowledge the system requires (Table 1). If the knowledge of the system about a particular task is not exactly the same as the knowledge of the user then the result can be different that the user expects. We can conclude that the higher level of abstraction we choose to communicate, the less data we have to transfer, and the less predictable the result is.

The evaluation of possible channels of communication (i.e. the device and the coding the channel uses) for transferring a particular syntax can include many aspects e.g., device costs, or the needed time and effort learning how to use the device etc. However, the main point of the efficiency of the transfer is the speed.

5. FURTHER RESEARCH

The research will go on by a comprehensive evaluation of the categories of the shape conceptualization techniques. We are going to use the abstract prototyping tool of Opiyo [4]. Based on the ranked categories, we will define the actual techniques to use. Since our aim is not to produce a fully functioning software prototype but only a proof of idea implementation, we will filter out several techniques, which is not necessary for the proofing and takes too much effort to implement.

REFERENCES

- [1] Rusák, Z., Horváth, I., Vergeest, J.S.M., Kuczogi, G., Jansson, J., "Discrete Domain Representation for Shape Conceptualization", Proceedings of the 4th International Conference of Engineering Design and Automation-EDA 2000
- [2] Shannon, C. E., Weaver, W., "The Mathematical Theory of Communication" The Board of Trustees of the University Illinois 1949
- [3] Horváth I., Rusák Z., Vergeest J. S. M., Kuczogi, G., "Vague Modeling for Conceptual Design", TMCE 2000 conference, pp. 131-144
- [4] Opiyo, E. Z., Horváth, and Vergeest, J. S. M., " Software Tools For Abstract Prototyping Of Design Support Tools", Proceedings of the 20th Computers and Information in Engineering (CIE) Conference: 2000 ASME DETC Conferences, September 10-13, 2000, Baltimore, Maryland, USA, in CD-ROM.

ВИЗНАЧЕННЯ ПЕРІОДУ ПАПІЛЯРНОГО ВІДБИТКА НА ЗОБРАЖЕННІ НА ОСНОВІ ПОБУДОВИ ФУНКЦІОНАЛІВ МАТЕМАТИЧНОГО СПОДІВАННЯ

Ростислав Косаревич, Ігор Ісаєв, Богдан Русин

Фізико-механічний інститут ім Г.В. Карпенка НАН України, 79601, м. Львів, вул.

Наукова 5,

тел.(0322) 65-45-30, e-mail: dep32@ah.ipm.lviv.ua

Представлено метод визначення періоду відбитка папілярних ліній на зображенні за допомогою побудови функціоналів математичного сподівання по випадковій реалізації. Розроблено новий підхід для визначення центра відбитка на зображенні.

1. ВСТУП

Сліди папілярних відбитків (ПВ) за своїм ідентифікаційним значенням займають перше місце серед біометричних характеристик особистості таких як обличчя, голос, підпис та ін. Візерунки, які утворюють папілярні лінії (ПЛ) нігтьових фаланг рук не змінюються протягом цілого життя. Якщо у однієї особи відібрати відбитки всіх десяти пальців, то вірогідність збігу із візерунками іншої людини буде рівна 1 з 64 млрд. Взявши до уваги загальну чисельність населення земної кулі, можна вважати, що збіг відбитків пальців двох людей практично неможливий [1].

На даний час широко поширені як автоматичні так і автоматизовані системи для ідентифікації особистості за допомогою візерунків папілярних ліній. Для роботи вони використовують або безпосереднє сканування поверхні нігтьових фаланг, або сканування відбитків візерунків папілярних ліній. В системах ідентифікації особистості, як правило, використовується другий спосіб оскільки він більш зручний. Результатом сканування є зображення розміром 512×512 , що має 256 градацій яскравості сірого кольору (Рис.1).

Для обробки зображень ПВ та ідентифікації особистості на їх основі, важливе значення має величина так званого періоду ПВ. Під ним розуміють суму величин ПЛ та міжпапілярного проміжку (МП) в одиницях виміру зображення. Ефективне покращення зображень ПВ, як правило, полягає у підвищенні контрасту між ПЛ та МП, а також в усуненні шумів на зображеннях. Таке покращення зводиться до застосування певного фільтра із заданими розмірами [2]. Ці розміри мають бути спільномірними з періодом ПВ, крім того

коефіцієнти фільтра та їх розташування залежать від діапазону в якому змінюються величини ПЛ та величини МП. Вважається, що товщина ПЛ є величина постійна, але при утворенні ПВ вона змінюється, оскільки поверхня пальця випукла, а слід переноситься на плоску поверхню. Крім того, при утворенні цифрового зображення слід квантується, що також вносить певні спотворення у величини ПЛ та МП, причому ці спотворення не є однаковим для різних ділянок ПВ. Коефіцієнти фільтра розташовуються так, щоб посилити ПЛ на зображенні та послабити МП. Таким чином для побудови фільтра необхідно знати діапазон зміни величини ПЛ та МП.

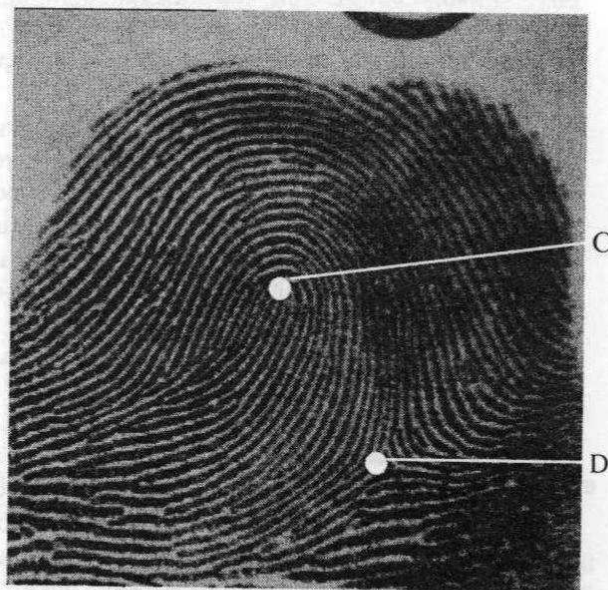


Рис.1. Особливі точки ПВ: С-центр; D-дельта.

В [3] запропонований метод оцінки періоду ПВ на основі профілів рядків та стовпців зображення. Цей метод дає можливість тільки приблизно оцінити межі зміни періоду ПВ. З метою більш точної оцінки пропонується для визначення періоду скористатися елементами теорії ПКВП.

Метод визначення діапазону зміни періоду ПВ за допомогою профілів має своїм недоліком те, що він не враховує розташування папілярних ліній на зображенні. У випадку, якщо ПЛ розташована під кутом відмінним від

90° величина періоду буде спотворена і це спотворення тим більше чим менший кут між прямою, що відповідає ряду або стовпцю зображення і папілярною лінією. Для більш точного визначення періоду ПВ та діапазону його зміни запропоновано новий метод, який полягає у скануванні зображення ПВ з його центра у декількох напрямках. На основі отриманих даних за допомогою методики ПКВП обчислюється період ПВ.

2. ЗНАХОДЖЕННЯ ЦЕНТРА ВІДБИТКА

2.1. Позначення і означення

Нехай D так званий "directional image" - поле напрямків напівтонового зображення G розміром $g \times g$, яке у випадку зображення ПВ містить ПЛ та МП. Напрямок у кожній точці представляє локальну орієнтацію ПЛ в цій точці. Напрямок $D_{(i,j)}$ в точці (i,j) зображення - це напрямок d , для якого S_d - сума різниць значень яскравості на зображенні вздовж напрямку d є мінімальною. Ця сума визначається як

$$S_d = \sum_{m=0}^{M-1} |G(i_m, j_m) - G(i, j)|, \text{ для } d=1, \dots, N,$$

де $G(i,j)$ та $G(i_m, j_m)$ - значення яскравості в точках (i,j) та (i_m, j_m) відповідно. Точка, яка відповідає (i_m, j_m) є m -ою точкою вздовж напрямку d від (i,j) , M - кількість точок вибрана для обчислення, N - кількість напрямків.

На основі поля напрямків D отримується поле напрямків блоків B (Рис.2) за допомогою знаходження домінуючого напрямку в блоці точок розміром $b_s \times b_s$. Це також зменшує розмірність з $(g \times g)$ до $(g/b_s \times g/b_s)$. Операції виконуються за допомогою обчислення гістограми напрямків H , яка формується на основі поля напрямків блоків B . Гістограма напрямків H - послідовність $\{H(0), \dots, H(N-1)\}$, де $H(j)$ - частота напрямку j в деякому околі.

В наших дослідженнях ми використовували $g=512$, $N=8$, $M=8$, $b_s=16$. Вибір величини g пов'язаний із вимогою якнайкращого представлення ПЛ на зображенні. Величина блоків b_s вибирається із припущення, що один блок містить принаймні одну ПЛ та МП. Величини кількості напрямків N та кількості точок M необхідних для обчислення напрямку пов'язані між собою. Більше число напрямків N вимагатиме збільшення кількості точок M . В наших експериментах, а також в інших дослідженнях відомих з літератури [4,5] вибір величини $N=8$ дав добрі результати. Більше N дає більш точніше представлення напрямків ПЛ на зображенні але також спричинює значне зростання обчислень для кожної точки зображення.

Гістограма напрямків H може бути наступних видів[4]:

а) пік гістограми зліва - коли напрямки менші або рівні ніж 90° домінуючі, іншими словами, сума частот напрямків які менші ніж 90° перевищує суму частот напрямків більших ніж 90°;

б) пік гістограми справа - коли напрямки більші або рівні ніж 90° домінуючі;

в) піки гістограми на краях - коли домінуючими є напрямки, які не належать діапазону 45°-135°.

Особливими точками зображення ПВ є точки центрів та точки дельт (Рис.1). За означенням точкою центру ПВ є точка, яка знаходиться на вершині найбільш внутрішньої ПЛ центрального потоку ПВ. Для точки центру ПВ область, що розташована зліва від неї буде мати гістограму, пік у якої зліва, а для області, що розташована справа - гістограма буде мати пік справа. Точкою дельти є точка в якій с три потоки ПЛ зустрічаються один з одним. Тому для такої точки додатково до критерію точки центру повинна виконуватись умова, що гістограма області, яка розташована нижче дельти, матиме піки на краях.

2.2. Знаходження особливих точок ПВ.

Оскільки поле напрямків блоків B зберігає характеристики напрямків ПЛ на багатоградацийному зображенні, то подальші дії по виділенню і знаходженню особливих точок ПВ доцільно проводити на ньому.

2.2.1. Виділення.

Якщо елементам B поставити у відповідність числа від 1 до 8 за наступним правилом: 0° , 180° → 1, 22°30' → 8, 45° → 7, 67°3' → 6, 90° → 5, 112°3' → 4, 135° → 3, 157°3' → 2, то як показали дослідження необхідними умовами для віднесення точки до розряду особливих є наступні вирази:

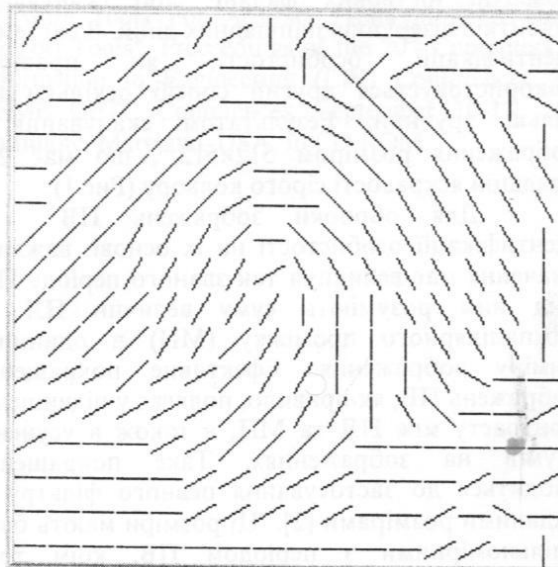


Рис.2. Зображення поля напрямків блоків B .

$$\begin{aligned}
 & B[j,i]=5 \wedge B[j,i+1] \in \{3,4\}, \\
 & B[j,i]=6 \wedge B[j,i+1] \in \{2,3,4,5\}, \\
 & B[j,i]=7 \wedge B[j,i+1] \in \{3,4,5\}, \\
 & B[j,i]=8 \wedge B[j,i+1]=2.
 \end{aligned} \quad (1)$$

При виконанні хоча б однієї умови (1) $B[j,i]$ відноситься до особливих точок ПВ. Про те спотворення вихідного зображення G можуть спричинити виявлення хибних особливих точок ПВ. Для того щоб цього не було запропоновано для кожної точки, що задовольняє вирази (1) перевіряти виконання наступної умови. Якщо позначити точки околу $B[j,i]$ розміром 3×3 через c_1, c_2, \dots, c_8 , то точка $B[j,i]$ не береться до уваги, якщо для неї не виконується умова $\sum_{i=1}^7 |c_i - c_{i+1}| \leq p$. Невиконання такої умови

означає, що окіл точки $B[j,i]$ є однорідним, а в самій точці є локальне збурення напрямку, що можливе при спотвореннях зображення G .

2.2.2. Кластеризація.

Наступним кроком знаходження особливих точок ПВ є кластеризація. Вона необхідна оскільки за означенням центра ПВ це є верхня точка найбільш внутрішньої ПЛ. Тому умовам (1) можуть задовольняти точки, які розташовані поряд на сусідніх ПЛ. Для зменшення їх кількості використовується алгоритм знаходження центрів кластерів. Далі особливими точками ПВ вважаються точки, що відповідають центрам виявлених кластерів. Для виявлення центрів кластерів використовується проста евристична процедура наведена у [6]. Її перевага в невеликих обчислювальних затратах, оскільки для виділення центрів кластерів достатньо однократного перегляду даних.

2.2.3. Відокремлення центра.

Після проведених операцій множина знайдених особливих точок ПВ містить точку, що відповідає центру ПВ та точки, які відповідають точкам дельт ПВ (якщо є). Для того щоб виділити єдину точку центра пропонується наступна процедура.

Спочатку розглядається окіл розміром 5×5 навколо кожної особливої точки. Цей окіл розбивається на чотири частини $R1, R2, R3, R4$ (Рис. 3).

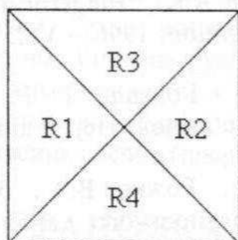


Рис. 3. Частини околу особливої точки.

Далі для кожної з частин $R1, R2, R3, R4$ формуються гістограми напрямків $H1, H2, H3, H4$ і в кожній гістограмі вибирається напрямок h_1, h_2, h_3, h_4 , що має максимальну частоту. Якщо для особливої точки ПВ величин h_1, h_2, h_3, h_4 задовольняють хоча б одній з наступних умов

$$\begin{aligned}
 & (h_1=4) \wedge (h_2 \in \{5,6\}) \wedge (h_3 \in \{1,2,7,8\}) \wedge (h_4 \in \{3,4,5,6,7\}), \\
 & (h_1=5) \wedge (h_2 \in \{4,5,6\}) \wedge (h_3 \in \{1,2,7,8\}) \wedge (h_4 \in \{3,4,5,6,7\}), \\
 & (h_1=6) \wedge (h_2 \in \{3,4,5\}) \wedge (h_3 \in \{1,2,7,8\}) \wedge (h_4 \in \{3,4,5,6,7\}), \\
 & (h_1=7) \wedge (h_2 \in \{4,5,6\}) \wedge (h_3 \in \{1,2,7,8\}) \wedge (h_4 \in \{3,4,5,6,7\}), \\
 & (h_1=8) \wedge (h_2 \in \{3,4,5,6\}) \wedge (h_3 \in \{1,2,7,8\}) \wedge (h_4 \in \{3,4,5,6,7\}),
 \end{aligned}$$

то вважаємо, що ця особлива точка – шуканий центр відбитка. Наведені умови відображають наведене в розділі 2 означення центра ПВ.

3. ФОРМУВАННЯ ВИПАДКОВОЇ РЕАЛІЗАЦІЇ

Після того як знайдено центр ПВ необхідно сформувати дані для обчислення періоду ПВ. Для цього з точки центра ПВ вздовж восьми ліній до перетину з краєм зображення G зчитуємо значення елементів (Рис. 4).

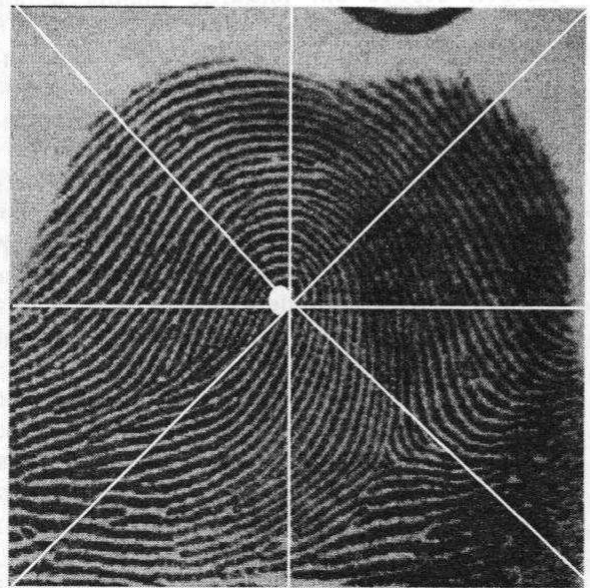


Рис. 4. Напрямки зчитування елементів зображення

Вибір напрямків зчитування пов'язаний із тим, що при правильній орієнтації ПВ проведені лінії в більшості випадків утворюють кути близькі до 90° із ПЛ на зображенні. Це дасть змогу більш точно оцінити період ПВ.

4. ЗНАХОДЖЕННЯ ПЕРІОДУ ПАПЛЯРНОГО ВІДБИТКА

Оскільки у випадковій реалізації, сформованій по ПВ, присутні періодичні коливання, модульовані випадковими збуреннями, найбільш адекватною моделлю для її дослідження і пошуку періоду, є модель у вигляді періодично корельованих випадкових

процесів (ПКВП). Ця модель є узагальненням детерміністичного і стаціонарного підходів до обробки сигналів. ПКВП визначаються як періодично нестационарні випадкові процеси з інваріантними відносно періоду корельованості математичним сподіванням і кореляційною функцією

$$m(t) = E\xi(t) = m(t+T),$$

$$b(t,u) = E\overset{\circ}{\xi}(t)\overset{\circ}{\xi}(t+u) = b(t+T,u),$$

$\overset{\circ}{\xi}(t) = \xi(t) - m(t)$. T - період корельованості процесу [7].

На основі теорії ПКВП у ФМІ НАН України були розроблені методи пошуку прихованих періодичностей, які попередньо ґрунтувалися на одній з перших історичних схем для виявлення періодичностей - схемі Бью-Балло. Ця схема має певні селективні властивості і при обробці реалізацій ПКВП. Після дослідження властивостей різних методів оцінки періоду, було показано, що метод максимальної правдоподібності дає асимптотично незміщену і асимптотично ефективну оцінку періоду корельованості ПКВП. Проте він вимагає знання детальної структури кореляційної функції. Були розроблені методи, які й програють по збіжності порівняно з методом максимальної правдоподібності, проте є більш простими по своїй реалізації, і менш вибагливими до апріорним відомостям про характеристики процесу. Ці методи базуються на властивостях функціоналів, побудованих подібно до виразів для когерентних і компонентних статистик ПКВП, з тою відмінністю, що замість істинного значення періоду, в них використовується деяка пробна величина.

Оскільки властивістю періодичності володіють математичне сподівання, дисперсія - на рівні перших моментів, кореляційна функція, спектральна густина і їх компоненти - на рівні других моментів, то побудова функціоналів для оцінювання періоду можлива по всіх цих характеристиках випадкового процесу. В даній роботі обмежимося дослідженням функціоналів, що побудовані на властивостях математичного сподівання.

Аналогічно до когерентного методу оцінювання ПКВП, побудуємо когерентний функціонал по математичному сподіванню для оцінювання періоду корельованості ПВ:

$$\hat{m}(t_0, \tilde{T}) = \frac{1}{2N+1} \sum_{n=-N}^N \xi(t_0 + n\tilde{T})$$

Когерентний функціонал при $\tilde{T} = kT$, k - ціле число, \tilde{T} - пробний період, приймає максимальні значення, що дозволяє уточнити межі, в яких знаходиться значення основного періоду ПЛ. Оскільки особливістю когерентного функціоналу є відбирання значень через

пробний період, то за його допомогою можливе знаходження тільки цілих значень.

Для точнішого визначення періоду використаємо схему, основу на інтегральних перетвореннях Фур'є. На рівні математичного сподівання розглянемо компонентний функціонал

$$\hat{m}_k(\tilde{T}) = \frac{1}{\theta} \int_{-\theta}^{\theta} \xi(t) \cos(\tilde{\Lambda}_k t) dt, \quad \tilde{\Lambda}_k = \frac{2k\pi}{\tilde{T}}$$

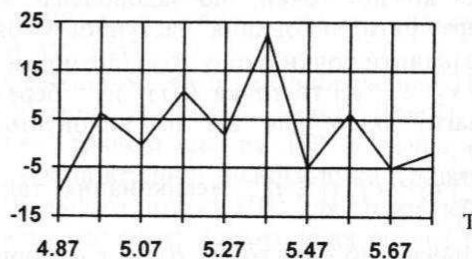


Рис. 5. Результат обчислення функціоналу 2-го компоненту математичного сподівання.

ВИСНОВКИ

Запропонований підхід до визначення особливих точок ПВ дозволяє ефективно проводити виділення центра відбитка.

По сформованій реалізації можливе оцінювання періоду ПВ за допомогою компонентних функціоналів математичного сподівання. Більш точне значення періоду досягається при використанні функціоналів по дисперсії та кореляційній функції.

ЛІТЕРАТУРА

1. Дубовий О.П., Лукашенко В.Я., Рибалко Я.В. та ін. "Криміналістичне дослідження слідів рук": Науково-практичний посібник. / За ред. Я.Ю. Кондратьєва. - К.: Атіка, 2000. - 152 с.: іл.
2. O'Gorman L., Nickerson J.V. "An approach to fingerprint filter design" // Pattern recognition.- 1989.- Vol. 22.- №. 1.- P. 29-38.
3. Косаревич Р.Я. "Метод визначення діапазону зміни періоду папілярного відбитка" / Матеріали XV відкритої науково-технічної конференції молодих науковців і спеціалістів ФМІ ім. Г.В. Карпенка НАН України "КМН - 2000", Львів, 2000. - С. 110-111.
4. Srinivasan V.S., Murthy N.N. "Detection of singular points in fingerprint images" // Pattern recognition.- 1992.- Vol. 25.- №. 2.- P. 139-153.
5. Karu K., Jain A.K., "Fingerprint classification" // Pattern Recognition, 1996. - Vol. 29. - №. 3. - P. 389-404.
6. Ту Дж., Гонсалес Р. Принципы распознавания образов: Пер с англ. - М.: Мир, 1978.- 411 с.
7. Драган Я.П., Рожков В.А., Яворский И.Н. "Методы вероятностного анализа ритмики океанологических процессов". - Л.: Гидрометеоздат, 1989. - 319с.

АЛГОРИТМ КЛАСИФІКАЦІЇ РОСЛИННИХ ОБ'ЄКТІВ З МОЖЛИВІСТЮ ШВИДКОЇ АДАПТАЦІЇ

Д. М. Піза, С. В. Морщавка

Запорізький державний технічний університет

м. Запоріжжя, вул. Жуковського 63

тел.: (0612)644662, факс.: (0612)642774 e-mail: svmorsh@zstu.edu.ua

У статті розглядаються питання розподілу об'єктів на два класи за допомогою алгоритму, який формує лінійні вирішальні межі по мінімуму метрики Махаланобіса. Вирішується задача скорочення часу навчання класифікатору. Пропонується алгоритм, який дозволяє, навіть при використанні великої кількості ознак, проводити навчання у режимі реального часу.

ВСТУП

У сучасних дослідженнях все частіше використовується багатоканальна зйомка у широкому оптичному діапазоні, від ультрафіолетової області до довгохвильової частини інфрачервоного діапазону. Спектральна інформація, що отримується при зйомці багатоканальними датчиками, часто служить для виділення одних об'єктів на фоні других. При цьому у якості ознак для такого розподілу служать коефіцієнти спектральної яскравості на різних довжинах хвиль, що відповідають різним діапазнам.

Прикладом таких багатоканальних систем може бути розпізнавання різноманітних об'єктів з космосу. З другого боку, широке використання флуоресцентної мікроскопії і спектроскопії для дослідження біологічних об'єктів можна також віднести до подібних задач. Ще однією важливою галуззю використання багатоканальної інформації про оптичні властивості об'єктів є сільське господарство. Така інформація може бути потрібна не тільки для визначення врожайності або засміченості посівів з космосу [1], але й для безпосереднього управління сільськогосподарською технікою в залежності від стану або виду рослини, яка оброблюється у цей час.

Така технологія була запропонована в ЗДТУ [2]. В ній, зокрема, коефіцієнти спектральної яскравості використовуються для розділення культурних рослин та бур'янів. Рішення про клас рослини, що оброблюється, використовується в подальшому для вибіркового знищення бур'янів (прополки) або часткового корисних рослин (проріджування).

1. ПОСТАНОВКА ЗАДАЧІ

Як показують дослідження, що були проведені

як у ході розробки означеної технології, так і іншими вченими [1, 3] існує декілька особливостей при розробці методів розпізнавання по багатоканальних даних.

По-перше, для отримання більшого обсягу вхідної інформації і, відповідно, більш якісного розпізнавання, необхідно використовувати як можна більше ознак. Наприклад, програмний комплекс MultiSpec, який був розроблений в університеті Пурдю (Індіана, США), використовує у якості вхідних даних знімки поверхні Землі, які зроблені літаковим спектрометром AVIRIS на 220 довжинах хвиль [3]. Нове покоління спектрометрів може дати ще більше інформації – до декількох тисяч спектральних інтервалів.

Друга особливість – все більша кількість задач потребує, щоб класифікація об'єктів проходила у режимі реального часу. Очевидно, що чим більше спектральних інтервалів (ознак) використовується, тим важче цього досягти.

Ще однією особливістю є необхідність використання алгоритмів з можливістю швидкого навчання або адаптивних алгоритмів для більш точного урахування змінних у часі оптичних властивостей об'єктів.

Всі ці особливості чітко просліджувалися і у ході вирішення задачі розпізнавання видів рослин для використання у новітніх сільськогосподарських технологіях. Як показали дослідження, що були проведені у 97-99 роках [4], для розпізнавання краще усього використовувати не одну-дві, а декілька ознак. Було також встановлено, що оптичні властивості рослин змінюються від поля до поля та залежать від періоду вегетації. Однак при цьому все рівно зберігаються відмінності культурних рослин від бур'янів. Таким чином, якщо проводити параметричну адаптацію класифікатора, то можна досягти практично стабільної якості розділення рослин.

Для класифікації виявилось можливим використовувати оптимальні алгоритми, що приймають рішення по максимуму правдоподібності. Така можливість є завдяки наявності нормального розподілу оптичних властивостей рослин [1]. Найбільш оптимальним по співвідношенню швидкості та якості розпізнавання було визначено алгоритм, який застосовується у лінійному дискримінантному аналізі. Він використовує

класифікацію по мінімуму метрики Махаланобіса при одній ковариційній матриці для обох класів (модель Фішера) [4]. Схема відповідного класифікатора зображена на рис.1.



Рис. 1. Схема класифікатора.

При цьому, якщо вважати вірогідності появи і штрафи за переплутування класів однаковими, то вирішальне правило виглядає як [5]:

$$\left(X_k - \frac{U_1 + U_2}{2} \right)^t V^{-1} (U_2 - U_1) \begin{matrix} < 0, & H_1 \\ > 0, & H_2 \end{matrix} \quad (1)$$

де: X_k – k -й тестовий вектор; U_i – середній вектор для класу i ($i=1,2$), якого отримано по навчальних вибірках; t – знак транспонування; H_i – рішення, яке приймається при дотриманні одного з відповідних нерівностей ($i=1,2$); V^{-1} – обернена ковариційна матриця, що відшукується за формулою:

$$V^{-1} = \left(\frac{V_1 + V_2}{2} \right)^{-1}, \quad (2)$$

где: V_i – ковариційна матриця для класу i ($i=1,2$), яка отримується по навчальних вибірках.

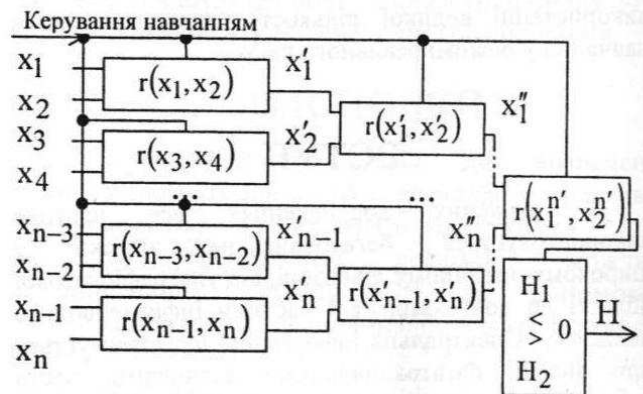
Такий алгоритм дозволяє реалізувати процедуру розпізнавання з мінімальним використанням обчислювальних ресурсів. Ще однією перевагою такого алгоритму, як і всіх параметричних, слід вважати достатньо просту процедуру його адаптації до змінних властивостей класів рослин. Для такого навчання достатньо тільки заново, або за допомогою оцінки, що ковзає, отримати нові вектор середніх та ковариційні матриці, що описують властивості класів рослин.

Однак при безпосередній реалізації навчання або адаптації класифікатора рослин по багатоканальних даних виникає одна суттєва проблема. Вона полягає у необхідності усякий раз обертати ковариційну матрицю, що входить у (1). При цьому у разі збільшення кількості спектральних інтервалів (ознак) – ситуація значно ускладнюється. Тільки кількість операцій множення та розділення, які виконуються найбільш довго, росте пропорційно п'ятому ступеню

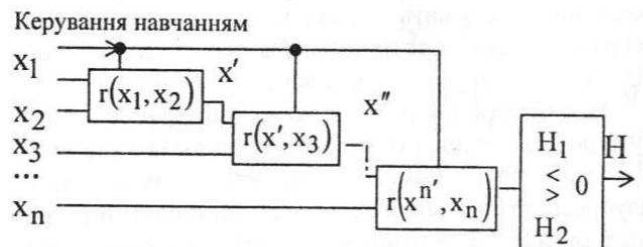
розмірності ковариційної матриці. Таке виходить, якщо знаходити мінори через приведення матриць до трикутного вигляду. Якщо ж використовувати прямий алгоритм для знаходження відповідних детермінантів, то кількість операцій зростає пропорційно факторіалу розмірності. Це значно знижує швидкість адаптації, і тому не дозволяє проводити навчання у реальному часі. Спосіб вирішення цієї проблеми пропонується у наступному розділі.

2. МЕТОДИ ВИРІШЕННЯ

Для зменшення витрат часу, необхідного для адаптації класифікатора, пропонується замінити традиційну схему з одноетапним пониженням розмірності інформаційного вектора (рис.1), на кількоетапну (рис.2).



а) паралельна схема класифікатора;



б) послідовна схема класифікатора.

Рис. 2. Схеми класифікаторів з поступовим зниженням розмірності.

При цьому вигреш буде досягтися за рахунок невеликої кількості операцій у комірках пониження розмірності, и того, що кількість таких комірок зростає набагато повільніше п'ятого ступеню розмірності.

Очевидно, що найбільший вигреш за швидкістю буде при використанні схеми з комірками, що мають два інформаційні входи и один керуючий (рис. 2). Для того щоб з'ясувати принцип функціонування таких комірок слід розглянути процедуру класифікації двовимірних векторів до одного з двох класів. Відповідна геометрична інтерпретація для цієї ситуації при застосуванні моделі Фішера наведена на рис.3.

Задача класифікації в такому випадку полягає у тому, щоб вирішити по яку сторону від межі

знаходиться поточна точка (x_{km}, x_{kh}) . Для того, щоб визначити це можна використати відстань r від цієї точки до межі, що вирішує.



Рис. 3. Геометрична інтерпретація класифікації в двовимірному просторі ознак.

Враховуючи, що межа лінійна, відстань від точки до неї можна знайти як:

$$r = \frac{1}{\sqrt{a^2 + b^2}} (a \cdot x_{km} + b \cdot x_{kh} + c), \quad (2)$$

де a, b, c – постійні коефіцієнти, які входять в рівняння лінійної межі. Їх можна знайти з рівняння межі, яке виводиться з (1). Так, рівнянням межі, що вирішує, буде:

$$\left(X_k - \frac{U_1 + U_2}{2} \right)^t V^{-1} (U_2 - U_1) = 0, \quad (3)$$

або, якщо розписати для двовимірного простору ознак:

$$\left[x_{km} - \frac{u_{1m} + u_{2m}}{2} \quad x_{kh} - \frac{u_{1h} + u_{2h}}{2} \right] \rightarrow \rightarrow \begin{bmatrix} v_{mm} & v_{hm} \\ v_{hm} & v_{hh} \end{bmatrix}^{-1} \begin{bmatrix} u_{2m} - u_{1m} \\ u_{2h} - u_{1h} \end{bmatrix} = 0, \quad (4)$$

де: (x_{km}, x_{kh}) – координати точки в двовимірному просторі ознак на виході m -го і h -го спектральних каналів; v_{mm} і v_{hh} – варіації ознак m -го і h -го каналів, відповідно; v_{mh} – коваріація ознак між m -м і h -м каналами; (u_{1m}, u_{1h}) і (u_{2m}, u_{2h}) – координати середніх векторів першого та другого класів, відповідно.

Для зручності розрахунків можна помножити ліву і праву частини цього рівняння на детермінант коваріаційної матриці, вважаючи, що він не нульовий. В протилежному випадку, якщо детермінант дорівнює нулю, то межа, що вирішує, вироджується у точку на прямій, уздовж якої розмістилися всі навчальні вибірки (одновимірна задача класифікації), і тоді формула (1) стає

непридатною для обчислень. Проводячи прості перетворення можна знайти коефіцієнти:

$$a = v_{hm} \cdot (u_{2h} - u_{1h}) - v_{hh} \cdot (u_{2m} - u_{1m}), \quad (5)$$

$$b = v_{hm} \cdot (u_{2m} - u_{1m}) - v_{mm} \cdot (u_{2h} - u_{1h}), \quad (6)$$

$$c = \frac{1}{2} \cdot (v_{hh} \cdot (u_{2m}^2 - u_{1m}^2) + \rightarrow \rightarrow + v_{mm} \cdot (u_{2h}^2 - u_{1h}^2)) - \rightarrow \rightarrow - v_{hm} \cdot (u_{2m} \cdot u_{2h} - u_{1m} \cdot u_{1h}) \quad (7)$$

На практиці, для прискорення класифікації поточної вибірки, зручніше користуватися не формулою (4), а більш простим співвідношенням:

$$r' = \frac{1}{|a|} (a \cdot x_{km} + b \cdot x_{kh} + c). \quad (8)$$

При цьому:

$$r'(x_{km}, x_{kh}) = K \cdot r(x_{km}, x_{kh}), \quad (9)$$

де K – коефіцієнт пропорційності. Можливість такого підходу пояснюється тим, що абсолютне значення величини r не грає великого значення. Для класифікації вибірки в такому просторі важливо знати її знак. При обчисленні відстані (8) з використанням коефіцієнтів (5)...(7), якщо її знак буде позитивним, то, згідно (1), рішення приймається на користь другого класу, та навпаки.

Пристрій, що реалізує (8), можна представити як комірку пониження розмірності від двовимірного простору ознак до одновимірного. Стан виходу такого пристрою у свою чергу можна вважати ознакою, що містить у собі об'єднану інформацію з двох вхідних каналів.

Таким чином, при використанні комірок пониження розмірності з двома інформаційними входами, можна запропонувати варіанти побудови класифікаторів: с паралельною структурою (рис. 2а), с послідовною структурою (рис. 2б) або їх комбінацію.

Паралельний і послідовний варіанти досліджувалися з метою перевірки їх роботоспроможності.

3. ЕКСПЕРИМЕНТАЛЬНІ ДОСЛІДЖЕННЯ

Для досліджень використовувалися данні, що були отримані у ході вивчення спектрів відбиття від різних культурних рослин і бур'янів. У сукупності даних були представлені наступні рослини: культурні – соняшник та кукурудза (по 52 вибірки даних), бур'яни – щиріця, лобода, молочай, мишій, осот, пирій (по 24 вибірки даних). В кожен набір даних входили відносні (с точністю до постійного множника) коефіцієнти відбиття від рослин або їх надземних частин. Такі коефіцієнти були отримані

на 55 довжинах хвиль від 300 до 850 нм з кроком 10 нм і смугою випромінювання 10 нм. З них безпосередньо використовувалися 32 з найбільш інформативних частин спектру – від 420 до 650 нм та від 720 до 790 нм включно.

Експеримент проводився наступним чином. Вибірки корисних рослин розділялися на дві частини за видами рослин. Відповідно, експеримент також поділявся на дві частини: навчання і розділення кукурудзи та всіх видів бур'янів на два класи; перенавчання і розділення соняшників та всіх видів бур'янів на два класи. Це відповідає переходу від поля з однією культурою до поля з другою. Очевидно, що задача переходу з поля на поле з одним і тим же видом рослин, але з декілька відмінними оптичними характеристиками буде вирішуватися ще простіше.

Безпосереднє навчання та перенавчання розподілу рослин на два класи проходило наступним чином. В кожній частині експерименту випадково обиралася одна вибірка, та, маючи апріорну інформацію про її клас, змінювалися відповідні середній вектор і коефіцієнти варіації та коваріації в усіх комірках послідовно по шляху розповсюдження сигналу. Після цього оцінювалася середня по усім вибіркам вірогідність вірного розпізнавання без зміни параметрів. Така послідовність процедур повторювалася багато разів. В результаті такого навчання добре видно підвищення середньої по двом класам вірогідності вірного розпізнавання до 85% (кроки 1...50 на рис. 4).

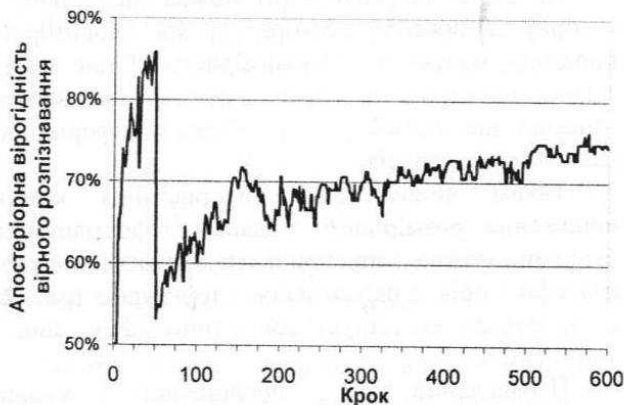


Рис. 4. Результати експериментів.

Після того, як вірогідність правильного розпізнавання переставала рости, набір вибірок замінювався на інший. При цьому вірогідність вірного розпізнавання різко зменшувалася (крок 51). І тільки після перенавчання вірогідність знов виростала до значень біля 75%. Слід відмітити, що, як і очікувалося, на початкове навчання потрібно набагато менше часу, ніж на перенавчання. В ході експериментів, для отримання середніх значень і коваріацій використовувалися згладжені оцінки.

Рис. 4 ілюструє результати, отриманні при дослідженні паралельної схеми побудови класифікатора. У випадку використання послідовної схеми досягти таких результатів виявилось неможливим. Проблема полягає в тому, що при

використанні поточної навчальної вибірки необхідно оцінювати середні значення та варіації і коваріації кожної комірки першого ступеню. Після цього знаходиться їх реакція на цю вибірку. І тільки потім оцінюються параметри у комірках наступного, по ходу розповсюдження сигналу, ступеню. Таким чином, для послідовної схеми шлях навчального сигналу зі входу на вихід виявляється довшим, ніж для паралельної схеми. Тому, внаслідок випадкового характеру навчальних вибірок, накопичуються помилки. В результаті такий класифікатор гірше навчається.

ВИСНОВКИ

Як показали проведені дослідження, запропонований алгоритм класифікації по багатоканальним даним дозволяє виконувати навчання і перенавчання у реальному часі. Кількість ознак при цьому може бути скільки завгодно великою. В той же час, зберігаються всі переваги алгоритмів, які класифікують по максимуму правдоподібності.

З двох запропонованих схем класифікаторів була обрана паралельна тому, що вона забезпечує більшу швидкість навчання. Послідовна схема простіше з точки зору програмної реалізації, але вона не забезпечує достатню точність і якість навчання при великій кількості ознак.

Ще однією важливою перевагою запропонованого алгоритму є можливість розпаралелення роботи комірок класифікатора, що визначається самою його структурою. Це дозволяє, навіть у випадку використання дуже великої кількості ознак, відносно просто забезпечити роботу у режимі реального часу.

ЛІТЕРАТУРА

1. Рачкулик В.И., Ситникова М.В. *Отражающие свойства и состояние растительного покрова.* -Л: Гидрометеоиздат, 1981.-287с.
2. Заявка на патент №94010291 (Україна). МКИ А01 В 39/18. *Способ прополки пропашных культур и устройство для его осуществления/ Пиза Д.М., Пиза А.Д., Татарчук И.Н.* -Заявл. 17.12.92 г.
3. Hoffbeck, Joseph P. and David A. Landgrebe, *Classification of Remote Sensing Images having High Spectral Resolution, Remote Sensing of Environment, Vol. 57, No. 3, pp 119-126, September 1996.*
4. Пиза Д.М., Морцавка С.В., Скоробогатов Ю.В. *Выбор эффективного метода распознавания растений по коэффициентам спектральной яркости.* «Радиоэлектроника, информатика, управление», №1, Запорожье: ЗГТУ, 1999р., с.85-88.
5. Айвазян С.А., Бухштабер В.М., Енюков И.С., Мешалкин Л.Д., *Прикладная статистика. Классификация и снижение размерности. Справочное издание.* М.: "Финансы и статистика", 1989,- 530 с.

ІНТЕЛЕКТУАЛІЗАЦІЯ ВІДЕОСИСТЕМ КОЛОВОГО ОГЛЯДУ

Боюн В. П., Довгань В. В.

Інститут кібернетики ім. В.М. Глушкова НАН України, Київ-187, пр. Академіка Глушкова, 40,
Факс 266-33-48, тел. 266-21-49, E-mail: vboyun@boyun.icfscst.kiev.ua

Запропоновано спосіб 2-х етапного перетворення відеопослідовності, отриманої з поворотної відеокамери, для підвищення "інтелектуальності" систем колового огляду. Перший етап – перетворення відеопослідовності у статичне панорамне зображення. Другий етап – виділення і реєстрація різнищового зображення між двома обертами відеокамери. При цьому значно покращуються умови роботи оператора, який здійснює контроль колового огляду, а також створюються умови для автоматичного виділення рухомих об'єктів або об'єктів, що змінюються, забезпечується значна компресія зображення. Алгоритм може служити основою створення інтелектуальних систем колового огляду.

ВСТУП

В нинішній час спостерігається широкий розвиток систем відеоспостереження (охоронні системи колового огляду, системи слідування, системи спеціального призначення і таке інше). З'являються системи відеоспостереження за виробничими процесами, об'єктами. Великі обсяги відеоінформації та складність аналізу рухомого на екрані зображення створюють дуже важкі умови для роботи оператора. Виходом з цього положення може бути підвищення інтелектуальності систем відспостереження, які можуть автоматично виділяти переміщення об'єктів, що відбуваються у зоні спостереження, або зміни геометричних розмірів, яскравості, кольору.

Найбільш розповсюдженими засобами компресії динамічних зображень, що враховують зміни у відеопослідовності, є стандарти MPEG-1, MPEG-2, які являють собою комбінацію ряду засобів компресії. В них використовуються засоби як просторової (всередині кадра), так і міжкадрової компресії. До перших відносяться засоби дискретно-косинусного перетворення і кодування по засобу RLE. Міжкадрова компресія базується на компенсації переміщення (оцінці вектору переміщення для блоку пікселів) і двунправленої інтерполяції кадру по двом сусіднім [1,2].

Ці засоби є універсальними і забезпечують достатньо високий ступінь компресії динамічних відеозображень. Однак вони вимагають виконання значних обсягів обчислень і складної апаратної підтримки для їхньої реалізації у реальному часі,

крім того вони орієнтовані на компресію рухомих зображень у системах реєстрації і передачі інформації і погано пристосовані для систем аналізу відеозображень в реальному часі.

Більш повне врахування умов колового огляду, особливостей таких систем і вимог, що пред'являються до них, дозволяють сподіватись на створення більш спеціалізованих і значно більш ефективних засобів компресії і виділення динаміки у відеопослідовності.

В доповіді розглянуто один з варіантів підвищення інтелектуальності систем колового огляду. Передумовами для вирішення поставленої задачі є:

- скорочення надлишковості у відеопослідовності;
- виключення з розгляду і аналізу постійного фону.

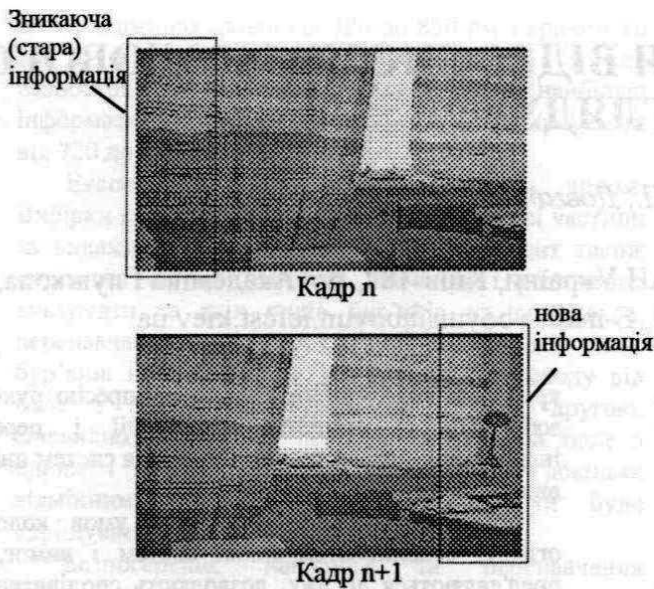
ПОБУДОВА ПАНОРАМНОГО ЗОБРАЖЕННЯ

Аналіз сусідніх кадрів відеопослідовності з камери колового огляду показує, що наступний кадр відеопослідовності містить значну частину попереднього, зсунуто в сторону, протилежну напрямку обертання камери. При цьому деяка частина зображення виходить з кадру і така ж частина нового зображення входить у кадр (див. мал. 1).

Якщо тепер з усієї відеопослідовності виділити тільки нові неповторювані частини зображень і "зклеїти" їх, то можна отримати широкоформатне панорамне статичне зображення одного повороту камери. Це панорамне зображення може бути представленим на екрані дисплею або у вигляді однієї широкоформатної смуги, або у вигляді 2-х, 4-х смуг по половині або чверті оберту відеокамери.

Основними проблемами отримання панорамного зображення є визначення границь зістиковки кадрів та згладження місця зістиковки для кращого візуального сприйняття.

При постійній швидкості обертання камери або при змінній швидкості із застосуванням датчика швидкості обертання зона нової частини зображення може бути визначена в залежності від швидкості обертання і від куту захвату відеокамери. Обсяг обчислювальної роботи для уточнення границі області не є великим.



Мал. 1. Кадри відеопослідовності при обертанні камери праворуч. Прямокутником виділена інформація, що з'являється у кадрі n+1 у порівнянні з попереднім кадром.

Ступінь стиснення при постійній швидкості визначається як відношення об'єму інформації у кадрі відеопослідовності до об'єму інформації у новій частині зображення або як співвідношення ширини кадру B і ширини нової частини зображення b .

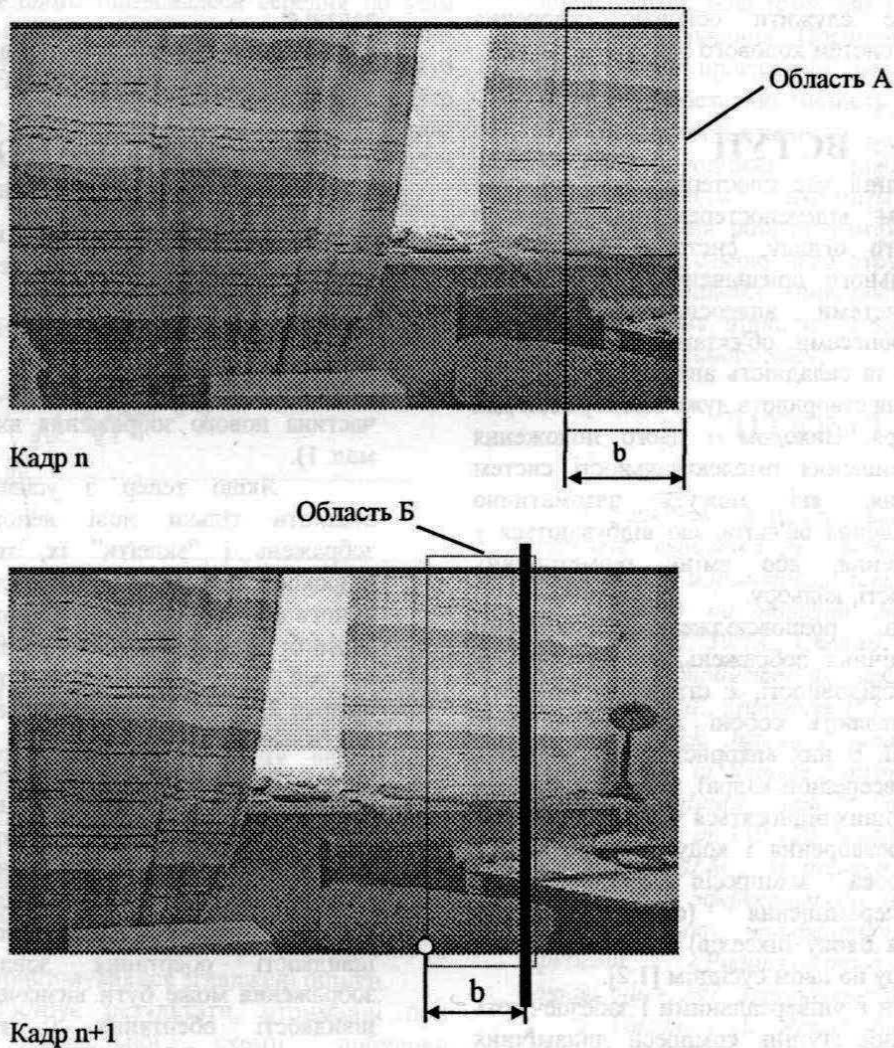
При змінній швидкості обертання відеокамери і відсутності датчика швидкості пошук границі попереднього кадру в наступному здійснюється шляхом сканування смуги краю зображення кадру n (область A) шириною b у зоні пошуку $B-b$ у кадрі $n+1$, починаючи з лівої сторони кадру.

Ступінь відповідності обраної смуги з кадру n в кадрі $n+1$ на кожному кроці сканування визначається, наприклад, по мінімуму функції [3.4]

$$F = \sum_{(i,j \in R)} |A(i, j) - B(i, j)|,$$

де $A(i, j)$ - значення яскравості елементів смуги краю зображення кадру n ;

$B(i, j)$ - значення яскравості елементів смуги, що обирається у зоні пошуку кадру $(n+1)$.



Мал. 2. Пошук границі появи нової інформації у кадрі в порівнянні з попереднім. Жирною лінією показана границя входження нової частини відеозображення.

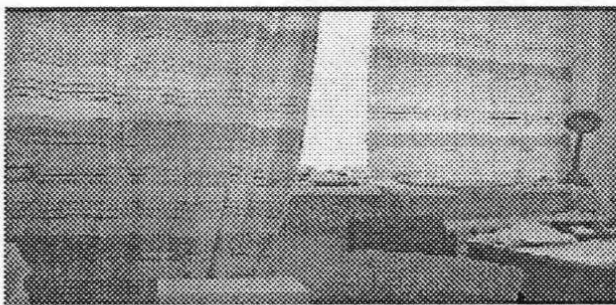
Мінімум функції F буде спостерігатися при визначенні ступеня подібності областей A та B (див. мал. 2).

Складність процедури пошуку краю зображення при нерівномірній швидкості обертання може бути визначена наступним чином.

Для однієї операції порівняння зі смугою $b \cdot h$ пікселів необхідно виконати $\approx 2bh$ операцій типу додавання. Якщо пошук здійснюється по зоні сканування шириною $(B-b)$, то загальна кількість операцій визначиться як $2(B-b) \cdot bh$.

Якщо визначити ширину контрольованої смуги краю зображення у відносних одиницях від ширини кадру як $b=k \cdot B$, де $k \leq 1$, тоді кількість операцій для реалізації процедури пошуку визначиться як $2kB^2bh(1-k)$.

Для поліпшення зорового сприймання панорамного зображення доцільно виконувати зістиковку нової частини зображення зі згладжуванням границі зістиковки [5] (зображення повинні плавно переходити друг в друга) – див. мал. 3, 4 а, б.



Мал. 3. Панорамне зображення, що є сумою кадрів, зображених на мал. 2.

Описаний спосіб панорамування відеопослідовності може бути використаний при проведенні зйомки з об'єкту, що рухається прямолінійно (наприклад, з поїзду, автомобіля). В результаті буде одержана широка панорама і значна компресія. Цей підхід може бути ефективно використаний при побудові систем віртуальної реальності [6].

Середній ступінь компресії відеопослідовності за рахунок панорамування сягає 10.

Представлення відеопослідовності у вигляді панорамного зображення значно полегшує умови контролю оператором зони спостереження.

Отримана послідовність широкоформатних відеозображень може додатково підлягати стисканню будь-яким алгоритмом компресії динамічних відеозображень [7], в тому числі і MPEG. При цьому забезпечується значно вищий ступінь компресії порівняно з компресією вихідної послідовності кадрів (ступінь компресії залежить від швидкості колового руху відеокамери, частоти кадрів зйомки, а також від динаміки зміни оточення).

Вхідне динамічне зображення може бути повністю відновлене з панорамного шляхом послідовного вирізування з панорамного зображення окремих областей. Для цього при складанні панорамного зображення слід зберігати ширину області, що додається до панорамного зображення від кадру до кадру (це необхідно тільки в випадку змінної кутової швидкості відеокамери). Динамічна послідовність відеокадрів може бути відновлена з панорамного зображення у темпі зйомки, або в довільному (уповільненому або прискореному) темпі.

ПОБУДОВА РІЗНИЦЕВОГО ЗОБРАЖЕННЯ

Наступним етапом зменшення надлишковості інформації, що представляється оператору, може бути вилучення фону з зображення і виділення нової інформації між обертами відеокамери. Це може бути реалізоване шляхом виділення різницевого зображення між панорамними зображеннями двох сусідніх обертів відеокамери.

На мал. 4 а, б показана панорама огляду відеокамери відповідно для першого та другого оберту камери. Як видно, зображення фону не підлягає змінам, розбіжність між цими двома зображеннями полягає в переміщенні предмету з положення A у положення B за час оберту камери. На підставі цих двох зображень побудоване третє, кожний піксел якого є різницею значень яскравості пікселів цих двох зображень (див. мал. 4 в). Якщо обмежити це зображення по порозі, та можна автоматично визначити не тільки наявність змін в обстановці, але і локалізувати місце, де відбуваються зміни. У результаті автоматичного визначення отримаємо дві області підвищеної уваги (на мал. 4 в вони виділені прямокутниками). Таким чином можливо робити автоматичну реєстрацію і сигналізацію наявності переміщення або локальної зміни яскравості (наприклад, при загорянні) в полі зору відеокамери колового огляду.

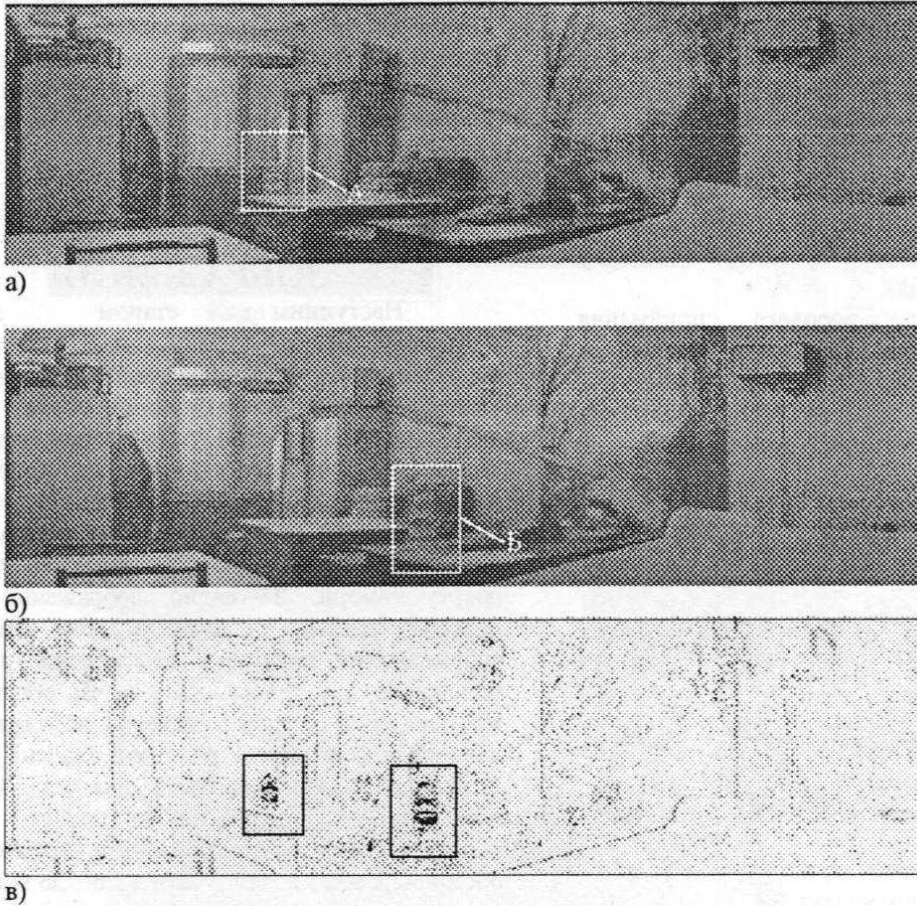
Похибки побудови панорамних зображень з ряду смуг, що складаються, границя між якими не завжди може бути визначена з точністю до піксела, а також наявність похибок від механіки системи колового огляду, призводять до появи деякого "шуму контурів" на різницевому зображенні. Однак такий шум у більшості випадків усувається високочастотною фільтрацією [5,8]. В той же час, зниклі, рухомі предмети або предмети що змінили геометричну форму, розміри, яскравість або колір будуть виділені практично повністю (якщо вони відрізняються від фону).

Використання різницевого панорамного зображення значно полегшує роботу оператора і скорочує обсяг інформації, що реєструється. В залежності від частоти і масштабності змін, що відбуваються ступінь компресії може досягати сотен

тисяч і більше разів. Якщо виконувалась постійна реєстрація (без фільтрації) різницевих зображень, то панорамні зображення, а при необхідності, і вхідна відеопослідовність можуть бути відновлені у реальному або довільному темпі.

Крім того, отримане різницеве зображення створює передумови для автоматичного викриття

змін, що відбуваються, тобто з'являється реальна можливість створення інтелектуальних систем колового огляду, що видають сигнал тривоги і реєструють виниклі зміни, або звертають увагу оператора на зміни, що мали місце.



Мал. 4. Приклад отримання різницевого панорамного зображення, побудованого з 160 кадрів відеопослідовності.
а) панорамне зображення, зняте на 1-ому оберті відеокамери
б) те ж на 2-ому оберті відеокамери
в) різниця зображень а) та б). Значення різниці посилені у 4 рази, зображення інвертовано.

ВИСНОВКИ

Запропонований спосіб панорамування відеопослідовності з камери колового огляду і отримання різницевого зображення, що виділяє зміни між двома обертами відеокамери, значно полегшує роботу оператора і скорочує обсяг інформації, що реєструється.

Крім того, забезпечується можливість створення інтелектуальних систем колового огляду, які автоматично виявляють зміни в коловій панорамі.

ЛІТЕРАТУРА

1. Зелов С. Цифровое кодирование изображений. Компьютер – Пресс №3,4 1997.
2. Сидоренко Ю. “MPEG”-Путь в цифровое завтра. Компьютерное обозрение. № 17(90) 1997.
3. В.В. Старовойтов. Локальные геометрические методы цифровой обработки и анализа изображений. – Минск: Институт технической кибернетики НАН Беларуси, 1997.-284 с.
4. Липанов А. В. , Путятин Е. П. Исследование алгоритмов обнаружения объектов на основе методов корреляционного распознавания и алгоритма параллельной нормализации. – Харьков: Радиоэлектроника и информатика №3, 1998.
5. Прэтт У. Цифровая обработка изображений. – М.: Мир. 1982.Т.2.
6. Хирозе М. Создание виртуальных миров на основе реальных изображений. Компьютеры+Программы №9(42), 1997.
7. М. Г. Георгиев, И.Г. Таранцев. Алгоритм сжатия последовательности изображений в системах визуального контроля редко меняющейся обстановки. – Новосибирск: Автометрия №2, 1999.
8. Абламейко С. В., Лагуновский Д. М. Обработка изображений: технология, методы, применение. – Минск: Институт технической кибернетики НАН Беларуси, 1999.-300 с.

ПОМІХОСТІЙКІ ШТРИХОВІ КОДИ

М. Ф. Козак, В. В. Ніколенко

УжДУ, каф. кібернетики та прикладної математики, Ужгород, вул. Підгірна, 46
тел. (03122)43325, jvtech@mail.uzhgorod.ua

ВСТУП

Штриховому кодуванню інформації, що надзвичайно широко використовується в країнах заходу, в нашій країні приділяється дуже мало уваги. Такий стан речей є незаслуженим – штрихові коди покликані полегшити та прискорити рутинну роботу людини в найрізноманітніших галузях людської діяльності. Широка сфера застосування штрихових кодів спричинила велику кількість найрізноманітніших типів штрихкодів. Нижче буде описано процес створення нового поміхостійкого типу штрихового коду Нем-8. Потреба в створенні цього коду викликана комп'ютеризацією контролю за переміщенням персоналу на підприємстві. Для цього кожного співробітника потрібно забезпечити пластиковою картою-перепусткою, на якій окрім фото та звичайної інформації про робітника, повинен був бути штриховий код для автоматичного зчитування інформації про робітника сканерами, що підключені до комп'ютера. Сформульовано вимоги до штрихового коду: простота та економічність у виготовленні, можливість кодувати до 8 чисел, і висока стійкість до помилок. Остання вимога найважливіша – при зчитуванні сканером штрихкоду, часто з'являються помилки. Причинами цього є ряд факторів: простота скануючих пристроїв, рух картки через сканер із прискоренням (робітник сам повинен протягувати її через сканер, а швидкість руху руки може бути нерівномірною), пошкодження штрихового коду у виді забруднення, подряпин, підтертостей. Помилки, що виникають внаслідок цього, є причинами того, що штриховий код не вдається розкодувати, або, що ще гірше, він розкодується неправильно, вони не дають змоги використовувати штрихові коди вже існуючих типів і виникає потреба в створенні нового поміхостійкого типу штрихкоду, основною особливістю якого є висока стійкість до помилок.

В ході роботи по створенню нового типу штрихового коду було вивчено та порівняно основні типи існуючих штрихкодів [3,4,5,6], на їх основі описано основні методи побудови нових типів штрихкодів, отримано математичний апарат, що дає змогу оцінювати стійкість коду до помилок, його надлишковість [1], та на основі цього апарату створено програмний пакет для дослідження стійкості до помилок різних типів штрихових кодів. Результатом цієї роботи став штриховий код Нем-8 основною відмінністю якого, від вже існуючих типів штрихових кодів, є його висока стійкість до помилок, яка була досягнута використанням в

штриховому кодуванні методів корекції помилок, що базуються на алгоритмі Хеммінга.

1. Способи побудови штрихових кодів

У штриховому кодуванні існує два способи графічного представлення бінарних значень '0' та '1'. Перший спосіб: значення '0' та '1' кодуються відповідно двома кольорами - білим та чорним. Наприклад: бітова послідовність 10110011101100011 буде мати наступне штрихове представлення (мал. 1). В цьому способі штрихи, що відповідають '0' та '1' мають однакову ширину. В разі якщо в бінарній послідовності йдуть одне за одним кілька однакових п значень '0' чи '1' їм буде відповідати білий чи чорний штрих п-кратної ширини.



Мал. 1

Другий спосіб представлення бінарних значень '0' та '1' у штриховому коді є спосіб, коли біти задані не різними кольорами, а різними значеннями ширини штрихів. Тобто є чотири атомарні графічні символи: два вузькі штрихи та два широкі, білого та чорного кольору. В такому штриховому коді білі та чорні штрихи чергуються, а значенням '0' та '1' відповідають відповідно широкі та вузькі штрихи. В цьому разі наведена вище бінарна послідовність буде мати вигляд (мал.2).

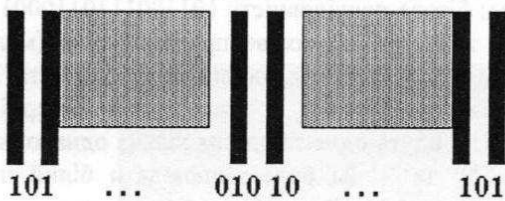


Мал. 2

Кожен із цих варіантів має як переваги, так і недоліки. В першому варіанті штриховий код буде коротшим, в наслідок того, що всі біти кодуються однаковими по ширині штрихами. По цій же причині в першому варіанті штриховий код бінарної послідовності зі сталим числом бітів буде мати сталий розмір, в той час як у другому варіанті розмір штрихового коду буде залежати від співвідношення нулів та одиниць. По цій причині перший варіант використовується якщо розмір інформації, що кодується, сталий, а другий варіант, коли кількість інформації, що кодується, змінна.

Розглянемо інші особливості побудови штрихових кодів, які також використовуються для

класифікації штрихових кодів. Однією з них є наявність чи відсутність контрольних штрихів (бітів). В випадку відсутності контрольних штрихів, при нерівномірній швидкості зчитування штрихкоду, є ймовірність, що цей код буде інтерпретовано неправильно. Щоб цьому запобігти, на початку та в кінці коду розміщується набір із принаймні двох контрольних штрихів. Після зчитування сканером цього коду, обчислювальна машина може судити про зміну швидкості сканування штрихового коду і відповідно корегувати процес декодування. Прикладом застосування контрольних штрихів може бути штриховий код типу EAN-13. В ньому контрольні штрихи наявні не тільки на початку та в кінці, а і в середині коду (мал. 3).



Мал.3

Ще однією особливістю при побудові штрихового коду є наявність чи відсутність контрольної суми. Для гарантування правильності декодування штрихового коду в деяких типах штрихового коду до інформації, що кодується, додається деяка контрольна сума, яка функціонально залежить від кодуваної інформації. Ця контрольна сума кодується в штриховий код разом з основною інформацією, а при декодуванні ЕОМ знову вираховує контрольну суму цього коду і порівнює із заданим. Зрозуміло, що неспівпадання цих двох контрольних сум свідчить про неправильно зісканований штрихкод. Поряд із контрольними числами, які кодуються разом з інформацією, нерідко в бітовій послідовності використовуються також біти контролю парності. Всі вищезазначені методи, покликані мінімізувати ймовірність неправильного інтерпретування штрихового коду, що сканується. В цьому штрихові коди чимось нагадують бітові послідовності сигналів, що передаються через цифрові канали зв'язку, тощо. Але штрихові коди мають одну дуже важливу відмінність, бітова послідовність, яка графічно представлена штрихкодом, не повинна містити велику кількість підряд розміщених однакових бітів, тому що графічне представлення такої послідовності, утвореної з одних тільки "0", перетвориться на тривалу темну смугу на штрихкоді, виміряти довжину якої для сканера буде складно. Тому, як правило, на штрихкоди накладається обмеження на кількість підряд розміщених однакових бітів, найчастіше це не більш ніж 3 однакові елементи. Все вище описане приводить до збільшення надлишковості інформації у штриховому коді, але в той самий час збільшує надійність коду. Надлишковість інформації присутня практично в

кожному типі штрихового коду в більшій чи меншій мірі, але у випадку штрихового кодування міру надлишковості можна умовно вважати мірою надійності (стійкості до помилок) штрихового коду.

2. Типи помилок, що виникають при штриховому кодуванні

В ході дослідження та створення штрихових кодів особливу увагу приділяють типам помилок, що виникають, та факторам, які стають причинами цих помилок. Найчастіше - це нерівномірна швидкість сканування штрихового коду, яка може бути наслідком багатьох причин: а) ручне сканування, дуже часто рука людини, що тримає сканер чи носій штрихкоду, рухається з прискоренням; б) траєкторія руху штрихкоду відносно сканера нелінійна; в) короткі зупинки штрихкоду під час сканування. Запобігти таким помилкам можна ще на стадії створення нового типу штрихового коду, ввівши в нього додатково контрольні біти-штрихи (мал.3), чи розмістивши поряд зі штрихкодом шкалу синхронізації сканера (мал. 4).



Мал.4

Нерівномірна швидкість сканування спричиняє помилкові ситуації, які досить важко виправити – порушуються пропорції в ширині штрихів і, як наслідок, в бінарному коді після сканування та декодування з'являються зайві біти, виявити котрі дуже складно.

Ще одним фактором, що спричиняє помилки при декодуванні штрихового коду, є забруднення та пошкодження штрихкоду. Такі проблеми виникають значно рідше ніж попередні, передусім, тому що з'являються, як правило, тільки при тривалому та частому використанні носія зі штрихкодом (посвідчення особи, бібліотечна справа, і т. п.). В таких випадках виникають три типи помилок: а) написи на штрихкоді, подряпини, перегини; б) плями; в) підтертості та іншого роду зменшення контрасту штрихового коду. В першому випадку пошкодження спричиняють заміну 1-2 бітів при декодування на протилежні. У другому випадку в бітовій послідовності з'являється тривала послідовність однакових бітів. В третьому випадку з'являється висока ймовірність неправильного декодування штрихового коду.

Усі ці три випадки формують специфічні задачі по розпізнаванню:

1. Локалізація та видалення написів із штрихової поверхні (мал.5).
2. Оцінка розмірів, локалізація та видалення плям (у разі якщо пляма не покриває повністю штрихкод). (мал.6).

3. Очистка штрихкоду та збільшення його контрастності (мал.7).



Мал.5



Мал. 6



Мал. 7

3. Нем-8

Нем-8 штриховий код створений для використання на посвідчені особи в межах великого підприємства. Цей штрихкод дозволяє кодувати 8-ми значні порядкові номери. Головною особливістю цього типу штрихкоду є використання методу Хеммінга для виявлення та корекції помилок. Серед існуючих типів штрихових кодів використовуються найрізноманітніші методи контролю за помилками, але методи корекції помилок, або не використовуються, або слабо реалізовані. Нижче описується процес створення штрихового коду Нем8.

Вимоги до штрихкоду, що були сформульовані до його створення:

Вимога 1. Мінімальна кількість символів, яку повинен містити штрихкод – 6. Можливо збільшити це число до 7 або 8, ввівши резервне або контрольне число.

Вимога 2. Алфавіт нашого штрихового коду складається тільки з 10 цифр 0,1, ...,9. Тому на кодування одного символу мінімально потрібно 4 біти $2^4=16$ ($2^3<10$) (насправді ця кількість буде дещо більшою, за рахунок уведення контрольних бітів та накладення деяких додаткових умов. Однією з таких умов буде наступна).

Вимога 3. Для запобігання злиття двох сусідніх бітових рядів, необхідно щоб кожен із таких рядів розпочинався з '0' та закінчувався '1'-цею.

Вимога 4. Велика кількість підряд розміщених однакових символів збільшує ймовірність помилки. Тому накладається ще одна умова: в бітовому ряді не повинно бути підряд більш ніж 3 однакових бітів.

Вимога 5. Як зазначалося у вимозі 2, в наслідок вимог 3-4, на кодування одного символу потрібно 6 біт.

Штриховий код Нем-8 забезпечено

властивістю самовиправлення помилок, яка базується на методі Хеммінга. Для цього необхідно з'ясувати, яку кількість бітів потрібно відвести під перевірочні. Виходячи з відомої формули $2^k > n+1$, де n - загальна кількість бітів, k - кількість перевірочних бітів, $m=n-k$ - кількість інформаційних бітів, отримуємо наступну таблицю:

Табл. 1.

K	N	m=k-n	K	N	m=k-n
1	1	0	5	31	26
2	3	1	6	63	57
3	7	4	7	127	120
4	15	11	8	255	247

Мінімальне число бітів у бітовому ряді 6, якщо під інформаційні відвести 26 біт, то можна було б кодувати $26/6=4$ символи. Така кількість є недостатньою, тому виділимо

Під інформаційні виділено 57 бітів (при 26 можна кодувати тільки $26/6=4$ символи), таким чином можна кодувати до $57/6=9$ символів, що є навіть надлишково для сформульованих вимог.

В таблиці 2 вказано всі можливі 6-ти значні бітові ряди, що задовольняють нашим умовам.

Табл. 2.

1	000101	5	001101	9	010111
2	000111	6	010001	10	011001
3	001001	7	010011	11	011011
4	001011	8	010101	12	011101

Цієї кількості бітових рядів достатньо, щоб кожному символі алфавіту штрихового коду було поставлено у відповідність по одному з них. Але розгляд матриці кодових відстаней цих бітових рядів, вказує на деякі недоліки набору. В різні виникнення однієї помилки в штриховому коді, вона буде виправлена за допомогою алгоритму Хеммінга. В разі виникнення двох помилок, метод Хеммінга перестане бути ефективним, в такій ситуації принаймні повинно бути виявлено наявність помилки. Перевірку на наявність подвійних помилок забезпечується вводом восьмого контрольного числа. Воно функціонально залежить від перших семи, і ця залежність вибирається так, щоб при зміні одного з перших семи чисел, восьме контрольне також обов'язково змінилось. Якщо при виникненні двох помилок, вони обидві знаходяться в одному бітовому ряді, то наявність такої помилки буде виявлено контрольним числом, але якщо ці помилки з'являться в різних бітових рядах, тобто будуть змінені два числа із семи, то можливі такі комбінації чисел, коли помилка не буде виявлена взагалі. Для запобігання таких помилок набір бітових рядів потрібно підібрати так, щоб матриця кодових відстаней цього набору не містила одиниць. Тоді при виникненні двох помилок у різних бітових рядах, ці новоутворені бітові ряди не будуть належати нашому наборові, що буде свідчити про помилку. В разі, якщо матриця кодових відстаней не містить одиниць,

то, для того щоб помилковий код не виявлявся, необхідно, щоб помилки виникли по дві в двох різних бітових рядах. Тобто вивід із матриці кодових відстаней одиниць, значно покращує стійкість коду до можливих помилок. Але для того щоб вивести з матриці кодових відстаней одиниць, потрібно скоротити набір бінарних рядів, але тоді цього набору буде замало щоб поставити один бітовий ряд у відповідність одному символу нашого алфавіту. З'являється необхідність у збільшення довжини бітових рядів з 6-ти до 7-ми бітів. При 7-ми бітових рядах можна кодувати $57/7=8$ символів. Це задовольняє вимогам. В таблиці 3 вказано всі можливі бітові ряди, що задовольняють сформульованим умовам.

Табл. 3.

1	0001001	9	0011011	17	0110001
2	0001011	10	0011101	18	0110011
3	0001101	11	0100011	19	0110101
4	0010001	12	0100101	20	0110111
5	0010011	13	0100111	21	0111001
6	0010101	14	0101001	22	0111011
7	0010111	15	0101011		
8	0011001	16	0101101		

Матриця кодових відстаней новоутворених бітових рядів також містить одиниці. Після відбору з 22-х бітових рядів десять таких, щоб їх матриця кодових відстаней не містила одиниць отримаємо таблицю 4.

Табл. 4.

0	0001001	4	0011101	8	0110011
1	0010001	5	0100111	9	0110101
2	0010111	6	0101011		
3	0011011	7	0101101		

При використанні цих бітових рядів Нем-8 повинен мати наступну структуру:

0xxxxx10xxxxx10xxxxx10xxxxx10xxxxx10xxxxx10xxxx
xxx10xxxxx1nnnnnn, де n – контрольні біти.

Ця структура містить поки що один невеликий недолік - останні шість бітів можуть бути причиною деяких помилок. Наприклад: а) вище ми накладали обмеження на кількість підряд розміщених однакових бітів, на комбінацію pnnnnn в даному випадку таке обмеження не поширюється; б) в разі, якщо останній байт чи байти будуть рівні "1", то при графічному представленні цього коду вони будуть втрачені, оскільки їм буде відповідати білий колір і вони зіллються з фоном штрихкоду. Щоб запобігти цим небажаним явищам, вводяться додаткові нульові біти в кінець коду. Кінцева структура коду:

0xxxxx10xxxxx10xxxxx10xxxxx10xxxxx10xxxx10xx
xxx10xxxxx101nn01nn01nn010

ВИСНОВКИ

Стійкість до помилок штрихкоду Нем-8 було перевірено за допомогою спеціально розробленої прикладної програми для ПК, яка

генерувала помилки та намагалася їх виявити чи виправити. Результати отримані за допомогою комп'ютера, підтвердили наші теоретичні розрахунки: при появі одного помилкового біта він виправляється завжди, при появі двох таких бітів будь-де в коді в 100% випадків фіксується наявність помилки та в 59% випадків виправляється. Помилки, які не виявляються, можуть з'явитись тільки при появі 4 помилкових бітів. Усе це свідчить про високу стійкість створеного штрихового коду.

Для порівняння штрихкоду Нем-8 з існуючими типами підрахуємо його ентропію. Нем-8 складається з 71 біта, кількість комбінацій чисел, які можна утворити 10 мільйонів (10^7), оскільки восьме число залежить тільки від перших семи, ми його не враховуємо. Кількість інформації, що несе в собі штрихкод Нем-8 рівна $I(a)=\log(10^7)$. В [3] наведена формула для знаходження величини надлишковості коду:

$$r = 1 - \frac{I(a)}{n \cdot \log L} \quad (1)$$

алфавіт двозначний $L=2$, кількість бітів $n=71$, тому

$$r = 1 - \frac{\log 10^7}{71 \cdot \log 2} = 0.67 \quad (2)$$

Коефіцієнт надлишковості штрихкоду EAN13:

$$r = 1 - \frac{\log 10^{12}}{95 \cdot \log 2} = 0.58 \quad (3)$$

Надлишковість в штриховому коді не є критичною, завдяки цій надлишковості коди забезпечуються більшою стійкістю до помилок, що при їх застосуванні більш важливо ніж надлишковість.

ЛІТЕРАТУРА

1. Шинон К. Работы по теории информации и кибернетике / Пер. с англ. под ред. Добрушина Р.Л., Лупанова О.Б. -М.: Издательство иностранной литературы, 1963. 830 с.
2. Котов П.А. Повышение достоверности передачи цифровой информации. -М.: Связь, 1966. 184 с.
3. Боховский В. А. Кибернетические системы в медицине. -М.: Наука, 1973. 328 с.
4. Cracking as an art, [Barcodes][Instant Access] / Lesson C(1). www.annon.penet.fi
5. Handbuch zum barcode-lesegerat / ABP-Programmierhandbuch fur Einbauscanner.
6. Barcodes FAQ. : www.azalea.com
7. Barcodes. : www.incom.ru/barcodes/s1.html

New Generation Video Surveillance System Based on Image Processing and FPGA Processor towards Autonomous System

A. Benkhalil, S. Ipson and W. Booth

Department of Electronic & Electrical Engineering
University of Bradford, Bradford, West Yorkshire, DB8 9EU, UK
akbenkha@bradford.ac.uk, +44-1274-234075

A stand-alone two-camera automatic security system is described. The system utilises a single Field Programmable Gate Array (FPGA) chip which performs a variety of video-rate arithmetic and logical operations on images, grabs and stores image frames and generates all necessary control signals. Results illustrating the highly successful operation of the system are included.

1. INTRODUCTION

Modern digital computing, semiconductor technologies and communication technologies have introduced a complete change of perspective in the design of the architectures of surveillance systems. Recently, several research groups have devoted a great deal of effort to devising prototype video surveillance systems based on robust methods of dynamic scene interpretation. Most of this work relies on complex and expensive hardware such as multiple DSP boards or on high-speed host computer platforms in order to achieve near real-time performance [1], [2], [3].

The research reported in this paper was directed at developing an image processing based surveillance system that does not need operator intervention, or a host computer or expensive hardware. The major aim of the research was to develop a stand-alone automatic security camera system (ASCS) capable of grabbing Comité Consultatif International des Radiocommunications (CCIR) format images, detecting motion and providing high resolution images of moving objects. A further aim was to take advantage of the high logic density of recently introduced FPGAs to implement motion detection and tracking algorithms and control functions in a single chip [4].

The paper is organized as follows. In section 2 an overview of the system is presented. In section 3 algorithms developed to detect motion, to update reference images and to carry out adaptive thresholding are described. Section 4 deals with the functionality of the FPGA including the implementation of the algorithm used for target classification. Section 5 presents some results to illustrate the performance of the system under different operating conditions and finally section 6 draws conclusions and indicates possible future developments.

2. SYSTEM OVERVIEW

A block diagram of our prototype ASCS, which employs two cameras working in a master slave combination, is shown in Fig. 1. The master camera points in a fixed direction and incorporates a 16mm lens to provide a field of view 22° wide in the horizontal direction. It provides greyscale images that are analyzed by the FPGA device every 40ms in order to detect target motion. The slave camera, which is mounted on a robotic arm for mobility, has a 150mm focal length lens to provide a high-resolution view of the target. It is moved to point at a target by the control logic hardware on the basis of information provided by the master camera. In the system currently implemented, the slave camera is driven to view one of the 100 sub-regions into which the field of view of the master camera is divided. The direction chosen is the one that is closest to the direction of the detected target.

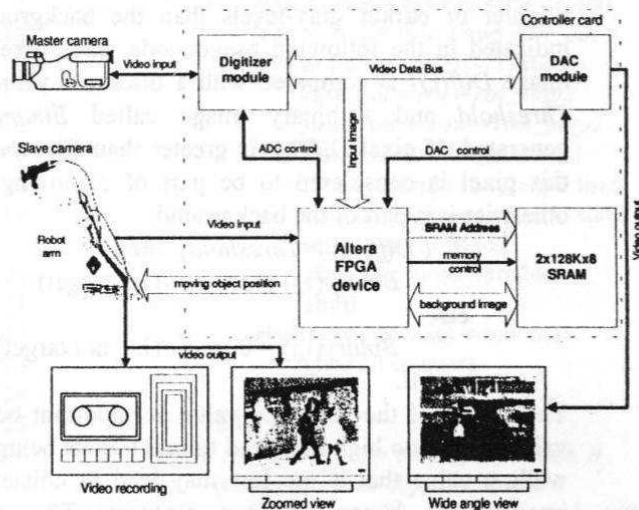


Fig. 1. Block diagram of the ASCS

The FPGA device shown in Fig. 1 is responsible for controlling the digitization and display of each incoming live-video image and the storage of a continually updated background image in the external SRAM. It also performs a variety of video-rate arithmetic and logic operations associated with moving target detection prior to transferring the co-ordinates of a detected target to the slave camera via the robot arm. A detailed description of the functionality of the FPGA is presented later in section 4. The digitizer module shown in Fig. 1 was implemented using a conventional flash Analogue-to-

Digital Converter (ADC) for digitization, a sync separator IC for timing control and a PLL (Phase Locked Loop) device to produce the system's 10MHz global clock from the 15.625 kHz CCIR line frequency signal. The Digital-to-Analogue Converter (DAC) module shown in Fig.1 uses an off-the-shelf DAC integrated circuit that automatically generates all the pedestal signals necessary to produce a standard video signal (CCIR). The system is designed to grab a stream of gray scale images with a resolution of 512(H) × 512(V) × 8 bits.

3. MOTION DETECTION

Since the system was intended to operate at video rates, a simple but efficient motion detection algorithm was chosen for implementation in the FPGA:

$$Diff(i,j) = | Live_im(i,j) - Bg(i,j) | \quad (1)$$

where $Live_im\{(i,j) \in R\}$ is the current image, $Bg\{(i,j) \in R\}$ is the background image and $R = \{(i,j): i=1 \dots M \text{ (512)}, j=1 \dots N \text{ (512)}\}$. This technique of subtracting the background image $Bg(i,j)$ from the input image $I(i,j)$ to produce the difference image $Diff(i,j)$ is called background differencing and it still forms the cornerstone of many of today's more complex systems [1], [2], [5]. The background image should, ideally, not contain moving objects and the system can use either a stored single live image or, more usefully, a continually updated stored image that is computed using a temporal filter as described at the end of this section. The absolute value of $Diff(i,j)$ is used in order to detect moving objects which may have brighter or darker gray-levels than the background. As indicated in the following pseudocode, each pixel of the image $Diff(i,j)$ is compared with a threshold value called *Threshold* and a binary image called $Binary(i,j)$ is generated. If pixel $Diff(i,j)$ is greater than *Threshold* then this pixel is considered to be part of a moving object, otherwise it is part of the background.

```

if ( Diff(i,j) > Threshold) then
    Binary(i,j) = 1    //(i.e. target)
else
    Binary(i,j) = 0    //(i.e. not target)
end if.

```

The choice of the threshold value is important because a value that is too high may lead to real targets being missed while a value that is too low may lead to noisier binary images and hence spurious features. The optimum threshold value also varies with the instantaneous illumination level and a suitable compromise has been achieved by continuously adapting the threshold value

$$Threshold = f \times \frac{\sum_{k=1}^{m \times n} Live_im(k)}{(m \times n)} \quad (2)$$

Here, $m \times n$ is the total number of pixels and f is a multiplying factor whose value can be set by the user. A value of $f = 1$ was found to be appropriate for the surveillance scenes used for testing so the threshold values actually used in these cases were the average gray levels of the live images. Fig. 2 demonstrates the detection of

moving features in a sequence of images captured using the master camera.

The background image is continually updated using the temporal low-pass filter [5] defined by:

$$Bg_{k+1}(i,j) = Bg_k(i,j) + G \times (Live_im_k(i,j) - Bg_k(i,j)) \quad (3)$$

Here k labels the current frame and G is a gain factor defined by the user. This feature causes the system to adjust the background automatically to take account of changes in the scene caused, for example, by changes in the illumination or by an object moving into the scene but then stopping. In either case the change of scene will become part of the updated background within a time-scale related to the value of G . Image sequence simulations carried out using MATLAB indicated that $G = 0.125$ is an appropriate value to choose. Then the background image takes about 20 frame periods to adjust to a step change in the scene, which enables new static features to become part of the background in a reasonably short time whilst ensuring that no moving objects are added to the background.

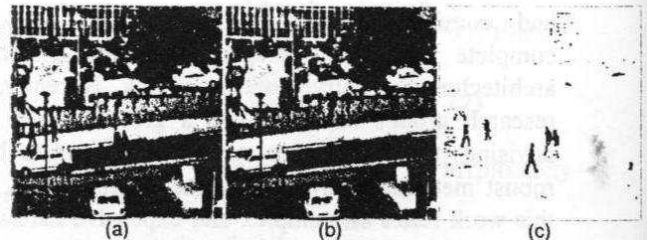


Fig. 2. The detection of moving objects. Here (a) is one of the images in a sequence of images, (b) is the current background image and (c) is the binary difference image after applying the current threshold.

4. FUNCTIONALITY OF THE FPGA

One FPGA chip forms the heart of the single board surveillance system. The functionality of this logic device is shown in Fig.3 by means of a schematic diagram. It can be seen that this chip performs a number of tasks, implemented as individual modules, in order to generate control signals for the slave camera from an analysis of incoming video images. The diagram also shows the data, address and control signal pathways connecting the different modules. Two of the FPGA's modules, frame-grabber control and memory control, provide all the control signals required by the external components in order to synchronize the image processing hardware to the CCIR composite video signal. These modules are implemented using 56 and 29 logic cells respectively, where 1 logic cell is equivalent to about 20 gates. The robot arm control module generates the signals required to move the slave camera and is implemented using only about 10 logic cells. Three image-processing modules are included in the FPGA. The first two of these, background image update and threshold selection, implement the algorithms described in the previous section within 17 and 68 logic cells respectively. Both algorithms require either a multiply or

divide operation (implemented by bit shifting) but the latter also requires an accumulator (implemented as a 20 bit adder) which accounts for the difference in the amount of resources used. The third image-processing module, target detection, is the most complex of all the internal modules and its implementation requires 286 logic cells. It decides whether a connected group of changing pixels is a target of interest or not. Because of its major influence on the workings of the system, the remainder of this section concentrates on describing the algorithm used to achieve the desired functionality.

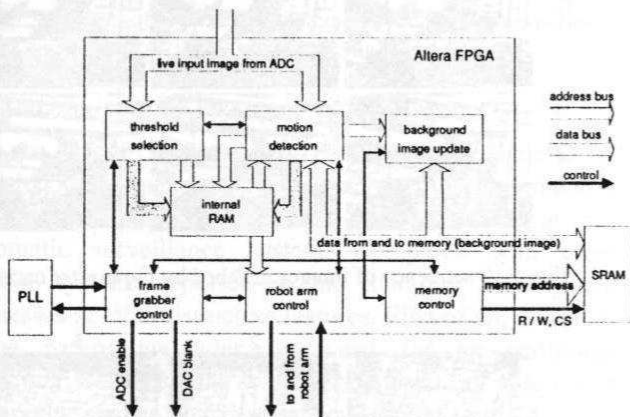


Fig. 3. Schematic diagram of the FPGA part of the overall system design.

The target detection module receives the live and background image data at video rates, pixel by pixel on a row by row basis, and applies differencing and thresholding as described earlier. The result of this process is illustrated in Fig. 4, which shows a small region cropped from Fig. 2(c). A potential target group of pixels, whose rate of change with time is faster than that set by the time constant of the temporal filter, is shown in black in Fig. 4 while background pixels are indicated in white. Moving along each row, a 4-bit counter col_count counts the number of successive pixels whose gray levels exceed *Threshold*. If the length of this run of black pixels is less than a user defined width limit col_udtp1 (appropriate to people) the counter is cleared. If however, this run length is greater than or equal to col_udtp1 another 4-bit counter, row_count checks the height of the identified object and compares it with a height limit row_udtp1 . The parameter row_count counts the number of successive rows in which runs of black pixels, starting at about the same column position, exceed the width limit. This situation is illustrated by position *b* in Fig. 4 for the case where col_udtp1 is chosen to be four. If row_count equals the height limit, then the co-ordinates of the target are saved in internal memory and also sent to the slave camera controller. This situation is illustrated by position *c* in Fig. 4 for the case where row_udtp1 is chosen to be 6.

The logic implemented in the target detection module is defined more precisely by the following pseudo-code. As indicated in this code, a section of the logic is repeated, the first time to detect large objects such as cars and the second time to detect smaller objects such as people. The

temporary variables and saved data indicated in the pseudocode are stored in the internal RAM of the FPGA.

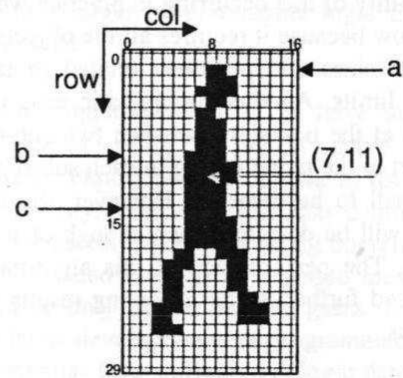


Fig. 4. A small region of Fig. 2(c) showing a detected object.

```
//within each of the 100 sub-regions do the following
1  Clear all counter variables to zero
2  for row =1 to end of current sub-region
3  for col =1 to end of current sub-region
4  if pixel at location [col,row] > Threshold then
5  col_count = col_count + 1
6  else
7  if col_count > or = to col_udtp2 then // (this
8  could be a large object i.e. car)
9  col_count = 0
10 if row_count zero then // (first row of object)
11 save_col = col - col_udtp2
12 save_row = row - row_udtp2
13 row_count = row_count + 1
14 else if row_count not zero then
15 // (subsequent rows of same object)
16 if save_row = row - 1 and save_col =
17 (col or col + 1 or col - 1) then
18 save_col = col - col_udtp2
19 save_row = row - row_udtp2
20 row_count = row_count + 1
21 if row_count = row_udtp2 then
22 send target co-ordinates to slave
23 camera controller and
24 clear all counter variables
25 endif
26 else if save_row not = row then
27 clear all counters
28 endif
29 endif
30 else if col_count > or = to col_udtp1 then //
31 (this is a smaller object i.e. person)
32 The previous lines 8 to 25 are duplicated with
33 col_udtp2 changed to col_udtp1, row_udtp2 changed
34 to row_udtp1.
35 endif
36 endif
37 end for
38 end for
```

This relatively simple algorithm detects only one potential target object within each sub-region and ignores any other potential target to the right of the first. This is of no concern in the current system because, as long as one target is detected, a view of that sub-region will be captured showing all targets that are present. However, it is conceivable that the presence of a number of objects,

smaller than the size threshold, to the left of a real target could prevent the detection of the real target. The probability of this occurring in practice would be expected to be low because it requires all the objects present to have specific sizes and positions related to target width and height limits. Another problematic case is when a target occurs at the boundary between two sub-regions, because the part of the target falling in each sub-region may then be too small to be detected. However, the target is moving and it will be detected when enough of it is within a sub-region. The performance of this algorithm in practice is discussed further in the following results and conclusions sections.

5. RESULTS

All the functions described in previous sections have been encapsulated within a single Altera EPF10K10LC84 device using the design and simulation tools provided by the Altera MAX+PLUS II development system. Two Altera library megafunctions, namely a fast 16-bit adder/subtractor and an allocator for the FPGA's internal RAM (EABs) were each used more than once in the design to shorten the design cycle time. It was found appropriate to create each of the six sub-designs shown in Fig. 4 as text files using Altera Hardware Description Language (AHDL). The complete design consists of about 1600 lines of AHDL source code. It uses most of the logic capacity and about 70% of the memory capacity of the specified device.

The complete system was tested by mounting the two cameras side by side, about 11m above ground level, and setting them to view a thoroughfare in our university campus that is used by both pedestrians and road vehicles. In order to illustrate the system performance, two synchronized frame grabbers were used to capture simultaneously image sequences from the master and slave cameras. Fig. 5 shows a typical short sequence of 12 images captured from the master camera over a 6s period on a rainy day. In this sequence one person (indicated by an enclosing black rectangle for ease of identification) can be seen walking down the sidewalk.

Fig. 6 shows higher resolution views of the moving person captured by the slave camera during the same period as it follows him. These images show the direction of the slave camera remaining fixed while the target moves through one of the 100 sub-regions, e.g. B1, B2 and B3, then changing to the next sub-region, e.g. C1, as the target progresses further. Because the slave field of view contains 512 by 512 pixels, there is sufficient resolution to recognize the person if the images are appropriately magnified. In this particular sequence, only one moving person is present and the slave camera accurately tracks the movement. When multiple moving targets are present, however, the system no longer tracks an individual target. It is designed to cause the slave camera to view them in order of their positions starting from the top left and moving from left to right and top to bottom through the list of target coordinates stored in the FPGA's internal memory. The list of targets is updated every frame period.

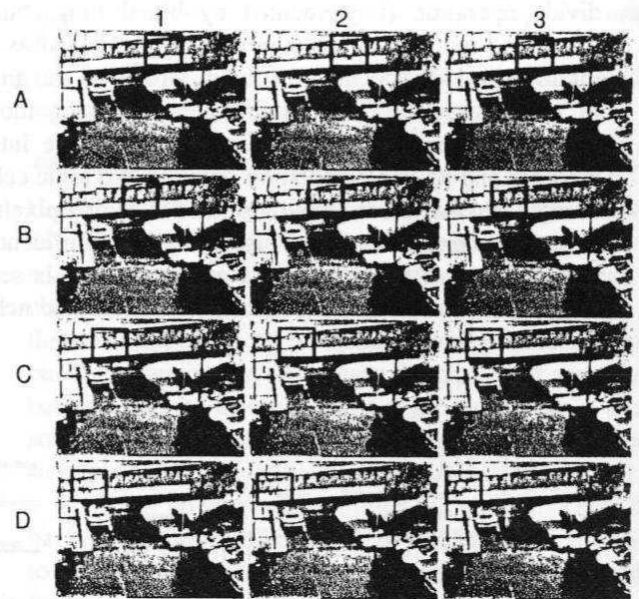


Fig. 5. A sequence of images grabbed by the master camera.

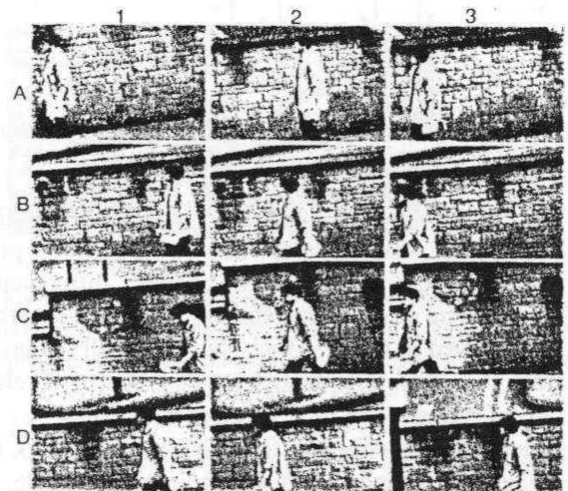


Fig. 6. A sequence of images grabbed by the slave camera.

Results on the detection and tracking of people recorded in one hour under different environmental conditions are shown in Fig. 7. The results demonstrate the ability of the system to work well in different situations. In bad atmospheric conditions, such as strong winds and heavy rain, the system generates a few false alarms but this is considered to be preferable to missing genuine targets. These false alarms arise from the movements of objects similar in size to people such as the branches of trees. On sunny days only a very small percentage of target people are missed, due entirely to their clothing being indistinguishable in gray level from the background. The possibility alluded to in the previous section of targets being missed due to the presence of other smaller objects was not observed in the trials and is much less important than the lack of contrast between some targets and the background.

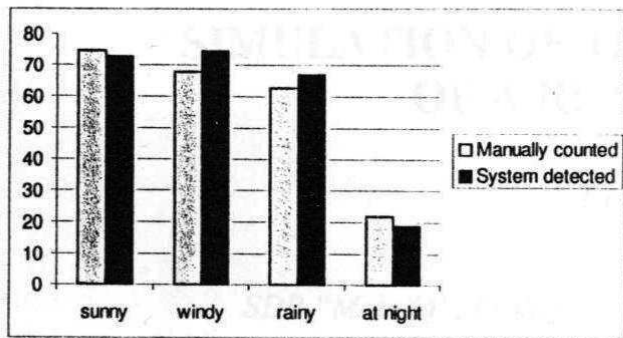


Fig. 7. The number of people detected over a one hour period under different environmental conditions.

6. CONCLUSIONS

All the major objectives of the research, namely the creation of a low cost stand-alone system incorporating a single FPGA controller, have been achieved in our automatic surveillance system. Compared with other similar systems reported in the literature, our system possesses several distinctive features. First of all, its stand-alone nature precludes the need for an additional computer. Secondly, the system operates at full video rates and grabs images and detects /tracks objects at 25 or 30 frames/s depending on the video standard used. Most of the alternative systems are significantly slower than this. Thirdly, the components on the board cost only about \$100 which is much less than all the other systems. Fourthly, unlike the other systems, it is specifically designed to provide several high-resolution views of all targets of interest within the scene. Like all of the systems mentioned in the introduction, our system is capable of working effectively under varying illumination and weather conditions.

Improvements to the current system could be made on two fronts. Firstly, the relatively slow speed action of the pan and tilt movement, which is currently provided by a robot arm, must be improved to take full advantage of the very fast detection and tracking performance of our system. This would involve developing a lighter and more accurate opto-mechanical means of changing the direction of the optical axis of the slave camera. By amending the control logic, the slave camera could then be directed to point at the center of each target of interest rather than at a sub-region. This would permit even higher optical magnification to be employed by the slave camera. The second improvement relates to image processing functionality. Although the current system works, it has less sophistication in terms of image segmentation and target classification than computer-based systems. Adding to the existing design an algorithm that fully segments the whole binary image in each frame period would enable all moving objects in the field of view to be located and improved target classification to be added. This would require extra internal memory to store one or two lines of run-length encoded live video data, temporary variables, arrays and the bounds of each segment comprising a group

of connected pixels. Additional logic resources would also be necessary. These modifications could be implemented in a larger programmable logic device at very little extra cost.

The successful implementation of a fully automatic surveillance system using a master slave camera combination is an important development in the area of vision-based security. The work has also demonstrated that a number of useful image-processing operations can be performed at video rates using a single inexpensive FPGA containing only 30000 useable gates. It is very clear that the latest developments in programmable logic devices, in particular their million gate logic density and their ability to perform several operations in parallel, afford new opportunities for researching and prototyping intelligent autonomous surveillance systems.

7. REFERENCES

- [1] G. L. Foresti, "A real-time system for video surveillance of unattended outdoor environments," *IEEE Trans. Circuits Sys. Video Technol.*, vol.8, pp. 697-704, Oct. 1998.
- [2] A. Iketani, A. Nagai, Y. Kuno, and Y. Shira, "Detecting persons on changing background," *IEEE 14th Int. Conf. on Pattern Recognition*, Queensland, Australia, 1998, pp. 74-76.
- [3] B. A. Boghossian, and S. A. Velastin, "Real-time motion detection of crowds in video signals," *Inst. Elec. Eng. Colloquium High Performance Architectures for Real-time Image Processing*, IEE, London, 1998, pp. 12/1-12/6.
- [4] D.W Trainor, D.J. Ridge, and R. Hamill, "Generic megafunction for high performance processing on FPGAs/PLDs," *IEEE Asilomar Conf. Signals, Systems & Computers*, Pacific Grove, CA, 1997, pp. 1047-1051.
- [5] A. Makarov, "Comparison of background extraction based intrusion detection algorithms," *Proc. IEEE Int. Conf. Image Processing*, Lausanne, Switzerland, 1996, vol. 1, pp. 521-524.

