IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 30 June 2023, accepted 21 July 2023, date of publication 25 July 2023, date of current version 3 August 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3298654

RESEARCH ARTICLE

Enhanced Artificial Vision for Visually Impaired Using Visual Implants

HOSSEIN MAHVASH MOHAMMADI[1], MOHAMMAD HADI EDRISI[1], AND YVON SAVARIA[2], (Fellow, IEEE)

[1]Computer Engineering Department, University of Isfahan, Isfahan 8174673441, Iran
[2]Electrical Engineering Department, Polytechnique du Montréal, Montreal, QC H3T 1J4, Canada

Corresponding author: Hossein Mahvash Mohammadi (h.mahvash@eng.ui.ac.ir)

This work was supported in part by Resmiq, and in part by the Natural Sciences and Engineering Research Council of Canada.

ABSTRACT Argus II is the most advanced retina implants approved by the US FDA and almost 350 visually impaired people are using it. This implant uses 60 microelectrodes implanted in the retina. The goal of this implant is to improve mobility and quality of life of its users. However, users' satisfaction is not very high due to the very low resolution of the phosphene images and features created by this device. This article proposes a system to improve the artificial vision created by visual implants. The proposed method extracts information about the people around the visually impaired person by using image processing and machine vision algorithms. This information includes the number of the people in the scene, whether they are known or unknown, their gender, estimated ages, facial emotions, and approximate distance from the visually impaired person. This information is extracted from the frames received by a camera mounted on the glasses of the user to generate signals that are fed into a visual stimulator. This information is shown to the user by a schematic vision created by some pre-trained patterns of phosphenes reflecting the information communicated to the user. The proposed system is validated with a simulated prosthetic vision comprising 150 microelectrodes that is compatible with the retina and visual cortex implants. A low-cost and energy efficient implementation of the proposed method executing on a Raspberry Pi 4 B at a frame rate of 4.5 frames/second shows the feasibility of using it in portable systems.

INDEX TERMS Argus II, artificial vision for visually impaired people, simulated prosthetic vision, scene understanding, visual prosthesis, visual implants, retina implant, visual cortex implant.

I. INTRODUCTION

According to statistics released by the World Health Organization (WHO) in August 2014 [1], 285 million people worldwide suffer from visual impairment, of which 39 million are completely blind.

In recent years, machine vision and artificial intelligence were used in visual implants and visually impaired mobility aids to help visually impaired people [2], [3], [5], [6], [7], [8], [9], [10]. Visual implants are intended to produce an artificial vision leading to some levels of functional vision restoration. Blind mobility aid systems can help to rehabilitate the visually impaired people for walking and avoiding obstacles.

The associate editor coordinating the review of this manuscript and approving it for publication was Mu-Yen Chen.

Visual prostheses can stimulate different regions on the visual pathways from the retina to the visual cortex to create an artificial vision. Based on where the stimulator is implanted, these prostheses can be divided into three groups: 1) retinal implants, 2) visual nerve pathway implants and 3) visual cortex stimulators.

Different kinds of visual implant have been proposed, but very few have passed all the clinical and technical trials allowing them to be commercialized. Many surveys on visual prostheses have been published [11], [12], [13], [14], [15], [16], [17].

ARGUS II proposed by the Second Sight Company is the only retinal prosthesis approved by the US FDA and it was reported that several hundred visual impaired people use it [18]. Second Sight is now merged into Vivani Inc. [19] and that corporation stopped producing Argus II. This implant

uses 60 microelectrodes implanted in the retina and can improve the quality of life of visually impaired people by making them experience light even if they were in the dark for many years. Visual acuity test accomplished with visually impaired people using Argus II shows that they can locate a big white square in a black screen or can recognize the direction or parallel white and black strips [20].

Orion®is a visual cortex implant for visually impaired people developed by the Second Sight Company which is merged into Vivani Inc. It targets a wider range of visually impaired people for whom the retinal implant does not work. This system is pending FDA approval for commercialization and it also uses 60 microelectrodes to be implanted in the visual cortex.

By contrast, Alpha IMS [21] includes an array of 1500 photodiodes implanted in the subretinal space. This system does not use a camera and an outside video processing unit to provide stimulus signals to be transmitted to the eye. Light sensors in this device are integrated on the subretinal space and it can restore some level of light perception. This device has been approved for commercial use in Europe [22].

Due to the limited number of microelectrodes of existing visual system stimulator, the artificial vision they permit has very low resolution. Many researchers have worked on improving the artificial vision created with low resolution implants by using image processing and machine vision algorithms.

This paper proposes a schematic vision system in which some patterns of phosphenes representing various significant information are defined. A variety of image processing and computer vision algorithms are applied to extract important information from the scene and people around the visually impaired person. This information is meant to be submitted to that person by various phosphene patterns. The proposed algorithms include face detection, face recognition, gender detection, age estimation, and emotion recognition. The main contribution of this work is to show how this information can be assembled to enhance the vision that low resolution visual implants allow to restore.

This paper is organized as follows. In the next section, an overview of previous works on artificial vision created by retinal or visual cortex implants is presented. The proposed method is detailed in the third section. The simulated prosthetic vision (SPV) is proposed in the fourth section. The fifth section reports results of experiments carried out to confirm the validity of the proposed system. Finally, the sixth section closes the paper with some discussion and conclusion.

II. PREVIOUS WORKS

Since the number of phosphenes in reported visual prosthesis is very limited, different research has been done to improve the artificial vision created by these prostheses.

Brindley and Lewin had already determined that the number of phosphenes required for an optimal reading of the letters must be 50 and this number for reading the handwriting must be greater than 600 [23].

Xia et al. [5] proposed a face semantic information transformation model to transform real faces into pixel faces using a face to pixel networks (F2Pnet). A pixel face database is also designed and a new training strategy for generative adversarial network is also proposed to solve the problem of semantic loss under limited number of pixels.

Steveninck et al. [3] proposed scene simplification methods in an indoor environment using deep learning-based surface edge detection and controlling the environmental complexity. They used a simulated prosthetic vision (SPV) with 26×26 phosphenes to provide sufficient information for mobility in a simple environment. Some methods exploit scene understanding and object detection to provide sufficient information for visually impaired people mobility in a simple environment [7], [24], [25], [26].

Avraham and Yitzhaky proposed an algorithm based on integral imaging to isolate objects based on their depth from the camera. They proposed an SPV system with 15×30 phosphenes. They also proposed different sizes and shapes of phosphenes obtained with different stimulation parameters [6].

Nicole et al. [2] proposed a scene simplification strategy using deep learning method to generate realistic predictions of SPV to improve scene understanding. They used an SPV with 32×32 phosphenes of different sizes.

Kartha et al. [27] proposed a disparity-based distance filter for Argus II. The application of the proposed filter on four experienced Argus II users was investigated. This filter improved the ability of the users for object localization, depth discrimination, orientation and size discrimination, and detection of people as well as their direction of motion.

A simulation of retinal prosthetic vision is proposed in which temporal aspects of the prosthesis such as persistence and perceptual fading of phosphenes and the electrode activation rate are considered [4].

Sanchez-Garcia et al. [24] proposed a schematic representation of indoor environments by phosphene image for object recognition and room identification tasks for indoor scenes. They extracted structural informative edges of the environment and silhouettes of segmented objects by using different convolutional neural networks.

Guo et al. [28]proposed two image processing strategies based on a salient object detection technique to improve recognition accuracy. Their method focuses on objects of interest and suppress the background clutter.

Parikh et al. [29] used feature extraction algorithms to identify the significant parts of an image and flashing phosphenes to attract the attention of a visually impaired person to a particular part of the visual field. Some authors reported improvements in object avoidance, reductions in head scanning and fast object localization with the use of cues [14].

Mohammadi et al. [30] presented an algorithm to identify the closest objects and to remove the background from a scene to create a sense of distance for the visually impaired person using specific phosphenes.

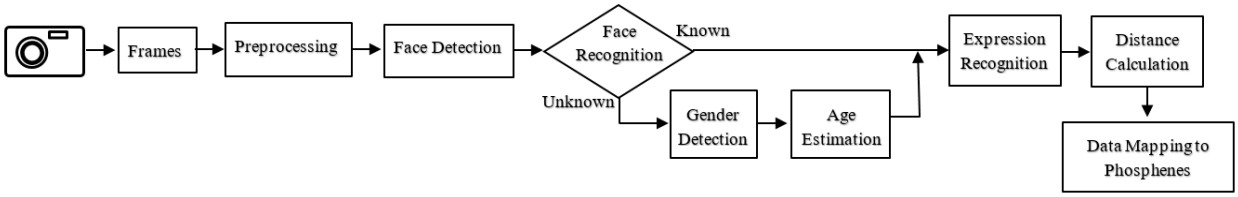

FIGURE 1. Block diagram of the proposed system.

III. PROPOSED METHOD

In the proposed method, a variety of image processing and computer vision algorithms are applied to extract important information from the scene and people around the visually impaired person. This information is meant to be submitted to that person by various patterns of phosphenes. While most of the previous works are focused on objects and environment around a visually impaired person, in this work we concentrate on people around that person.

The algorithms included in the proposed system comprise face detection, face recognition, gender detection, age estimation, and emotion recognition. For each feature, an appropriate algorithm was selected based on its accuracy and time complexity to enable affordable real-time implementations in an autonomous portable system. Thus, the main contribution of this work is not about proposing new algorithms for these functionalities but rather showing how they can be assembled to enhance the vision that low resolution visual implants allow to restore.

A block diagram summarizing the main steps of the proposed system is shown in Fig. 1. This system is assumed to comprise a camera mounted in the glasses of a visually impaired person to capture video frames from the scene in front of that person. In the first step, a face detection algorithm using Haar features and the Adaboost classification method extracts a square area around all faces in each frame [31].

In the second step, a face recognition algorithm is used to identify each face using the Fisherfaces algorithm [32]. Each image is preprocessed and adjusted prior to the face recognition procedure.

If a detected face is not recognized, a gender detection method using the Eigenfaces algorithm [33] is applied. An age estimation algorithm is applied in the fourth step, where a trained neural network is used to estimate the age group of each unrecognized faces.

In the fifth step, an expression recognition method using Local Binary Patterns (LBP) features [34] and Linear Discriminant Analysis (LDA) classification is applied. After this step, the results are transmitted to the visual implant using a phosphene map platform. The information extracted in the previous steps are shown in this phosphene map into different categories. The rest of this section provides more details on each phase and report on their respective performances.

A. PHASE 1: FACE DETECTION

The face detection algorithm presented by Viola and Jones [35], known as the Haar Classifier, is the first face detection method working based on reinforcement learning. It has been widely used in digital cameras and image processing software in practical applications. Due to its accuracy and real-time processing speed, the Haar classifier is used in the proposed system. In the Viola-Jones method, Haar features combined with the Adaboost learning method creates a strong and secure classifier that can quickly detect the faces in an image.

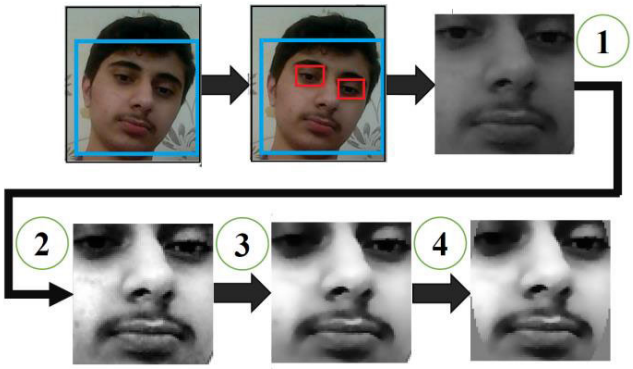

FIGURE 2. Four preprocessing phases performed on an image.

B. PHASE II: FACE RECOGNITION

Face recognition is the process of identifying specific individuals in an image and matching their faces with a library of known faces. Face recognition methods are very sensitive to the image conditions such as lighting, shadows, direction, and expression of the face. The various steps of preprocessing are applied to mitigate the challenges of these situations and to improve the accuracy of the face recognition method. In this study, the Fisherfaces algorithm is used for face recognition in three stages of preprocessing, training and recognition.

The preprocessing step includes the following operations:

1. **Geometrical transform and clipping**: Eye detection is used to detect the eyes positions and then a rotation is applied to align the eyes positions in a horizontal direction. The size of the bounding box of the face is also scaled to a predefined size and the forehead, chin, ears and background are also removed.
2. **Histogram equalization**: histogram equalization is applied on the left and right sides of the face independently to standardize the contrast and brightness of the input images.
3. **Smoothing:** A Gaussian filter is applied to reduce the image noise.

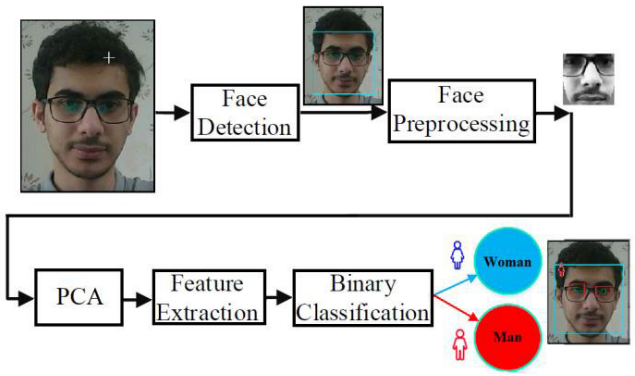

FIGURE 3. The basic framework of the gender detection system.

4. **Elliptical mask:** the image background out of an elliptical mask is eliminated.

Fig. 2 shows the four phases of preprocessing performed on a sample image.

In the training phase, preprocessed images of faces and their labels are stored in a library, and the training algorithm is performed using the Fisherfaces algorithm. To improve accuracy, the training set includes images captured under various lighting conditions, directions, and face expressions.

In the face recognition phase, detected faces are compared with the training datasets using the maximum likelihood method. For any given input, the system will select the label of the most similar image in the training data set. If the likelihood of the input image is less than a threshold, the input image is considered unidentified.

C. PHASE III: GENDER DETECTION

Gender detection is one of the challenging issues in machine vision whereas a human can easily recognize the gender with a 95% accuracy [36].

Generally, gender detection methods use face detection, image preprocessing, feature extraction and classification. In the proposed system, the eigenface method is used for gender detection. The basic framework of this method is shown in Fig. 3.

After face detection and preprocessing, a data size reduction technique using PCA (Principal Component Analysis) is applied in order to reduce the computational complexity of the method. In PCA all training and input images are mapped into their eigenspace comprised of eigenfaces which are based on a specific group of images. In the feature extraction phase, descriptors providing very distinct features of an image are selected. In the last step, a supervised learning classifier is used. The classifier uses limited number of eigenfaces to find the nearest neighbor between the input images and the training samples to classify them into two classes of man or woman.

D. PHASE IV: PROPOSED ALGORITHM FOR AGE ESTIMATION

The age estimation algorithm in the proposed system uses a Conditional Probability Neural Network (CPNN) [37].

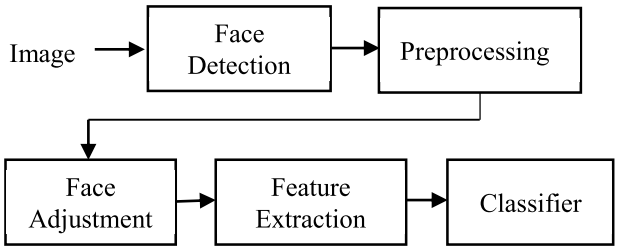

FIGURE 4. Block diagram of the facial expression recognition system.

We used two types of extracted features from the image as the feature vector. A statistical output assigned to each age label and age estimation is done using a multilayer feedforward neural network.

We use a three-layer neural network of CPNN network [37] to approximate p(x|y), the degree to which the y label of an input is observed given the input age vector of x.

We propose to use Active Appearance Models (AAMs) to obtain global features and the brightness of the skin for local features to present a hybrid model. These two types of features are independent of the race, lifestyle, and geographical conditions; therefore, they are more reliable than other features.

While CPNN [37] uses a feature vector of size 200, the proposed method applies PCA to reduce the dimension of combined feature vectors to 160 in which the lengths of local and global feature vectors are 24 and 136 respectively. The method also uses a Multilayer Perceptron (MLP) neural network with 3 layers and a total of 64 neurons in the hidden layer, sigmoid activation functions, and linear activation functions in output layer with a variance of 4. After training the network for an unseen input sample, a statistical output is obtained. Using the average estimate of the closest likelihood, a maximum value would be found for the output that can be used to estimate the age group. The method estimates the actual age with the mean absolute error (MAE) of 3.54 years.

E. PHASE V: FACIAL EXPRESSION RECOGNITION

Facial expression recognition systems (FER) recognize the mood/state of a face, namely, happiness, sadness, fear, disgust, anger, surprise, etc.

Two common strategies for extracting facial features are geometry-based methods and appearance-base methods. Valstar et al. [38] showed that, in practice, geometry-based methods work equally well or better than appearance-based methods, but they usually require accurate and reliable detection and tracking of face features that are not applicable in all cases.

Facial expression recognition consists of five steps, namely, face detection, image preprocessing, face adjustment to new coordinates, feature extraction (e.g., LBP), and classification as shown in Fig. 4.

Feature extraction of local binary patterns (LBP) was proposed for texture analysis and it has been gradually used

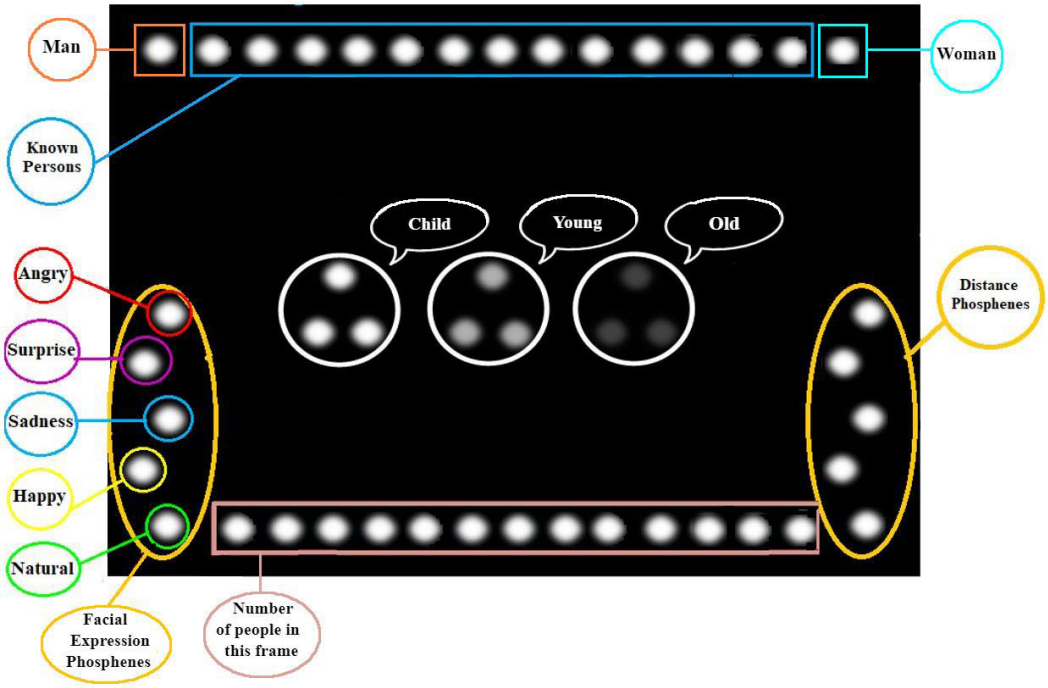

FIGURE 5. Proposed simulated prosthetic vision (SPV).

in other applications such as emotion recognition [39]. The most important characteristics of LBP are its resistance to luminance changes and lower computational complexity. Therefore, the original LBP [40] feature extraction method and the LDA classification for face expressions were used in the proposed system. Kumari et al. [41] were conducted a comparison between different classification and LBP feature extraction methods.

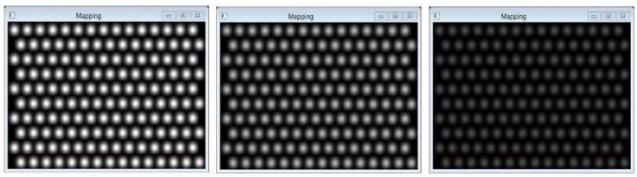

FIGURE 6. Simulated phosphenes on a hexagonal grid with three active levels of brightness.

IV. SIMULATED PROSTHETIC VISION

This paper proposes a simulation of the artificial vision in which the information synthesized by the system introduced in Fig. 1 is shown to the visually impaired user using a low resolution phosphene image. The simulation introduces an SPV in which some patterns of phosphenes representing various information are defined. The system introduced in the previous section is also validated with the proposed SPV to create a mode of artificial vision for users.

As discussed earlier, the number of phosphenes in visual implants is limited. Argus II and Orion use 60 microelectrodes. Pixium Vision Company uses 142 microelectrodes in PRIMA retinal implant, and 150 microelectrodes in IRIS III and the company is going to use 260 microelectrodes in its next generation retinal implant. A comparison between different retinal implants and number of microelectrodes used in each prosthesis is reported by Sohmyung Ha et al. [13].

The proposed simulator assumes that an implant can stimulate 150 phosphenes organized according to a hexagonal 10×15 mesh offering three levels of brightness. This number of phosphene is larger than those reported in existing artificial vision stimulators, but as the need for higher resolution is obvious, this research explores what could be done with next generation visual implants. If a greater number of microelectrodes is applicable, more information can be shown and new patterns for a higher resolution of phosphenes can be redefined accordingly.

Different number of grayscales (3, 5, 10 and 12) were reported in both retinal and cortical prostheses [42]. In the proposed SPV three levels of phosphene gray scales are used. Fig. 6 shows the SPV with three active gray levels when all phosphenes are turned on at a given level. Fig. 5 shows the SPV in which the outputs of the face recognition, gender detection, emotion recognition, and age group estimation. In addition, the estimated distance of the person and the number of people in the frame are mapped onto separate groups of phosphenes.

As shown in Fig. 5, a pattern of three phosphenes is used to represent a detected face. This pattern moves around the frame with respect to the location of the face in the video frame.

TABLE 1. A comparison between the proposed system with the most recent works.

Authors	Year	Number of Graylevel	Size of SPV	Features proposed	Speed FPS	System platform
Proposed system	2022	3	10 × 15	Face recognition, gender detection, age estimation, Expression recognition, depth estimation	6 4.5	Core™ i7-2640MHz Raspberry Pi 4 B
Xia et al. [5]	2022	N/A	25 × 23	Facial expression recognition	15	Raspberry Pi 4 B
Van et al. [3]	2022	2	26 × 26	Scene simplification in an indoor environment	18.3	N/A
Beyeler et al. [7]	2022	VR head-mounted display		Visual augmentation, visual orientation mobility	N/A	N/A
Avraham et al. [6]	2021	4	15 × 30	Objects isolation based on their depth	N/A	N/A
Han et al. [2]	2021	256	8 × 8, 16 × 16, 32 × 32	People and car identification	N/A	N/A
Kartha et al. [27]	2020	This system was tested on Argus II Implant		Object localization, depth discrimination, orientation and size discrimination, and people detection and direction of motion	N/A	Processor with a Xilinx Zynq-7000 SoCs.
Sanchez-G. et al. [24]	2020	8	32 × 32	Object recognition and room identification tasks	N/A	N/A
Lozano et al. [25]	2020	N/A	15 × 15	Semantic segmentation and object detection	14, 20	N/A
Jiang et al. [26]	2020	VR head-mounted display		Complex scene perception, object detection	N/A	Core™ i7-9700k 3.60GHz, 32GB
Guo et al. [27]	2018	256	32 × 32	Salient object detection	N/A	Intel Core 2.8GHz with 2GB RAM
Wang et al. [43]	2014	8	24 × 24, 32 × 32	Region-of-interest magnification and face detection	1-15	Core™ i3–550 CPU @3.20 GHz
Parikh et al. [29]	2013	8	6 × 10	Important objects detection,	N/A	N/A
Lui et al. [44]	2012	2	25 × 25	Preserving salient features such as edges.	25	Intel i5, Kinect
Mahvash et al. [30]	2012	2	30 × 32	Background removal and sense of distance	0.5	Core i7-860

Different levels of brightness on this pattern are used for showing age groups (children, young, and old people) to which the detected faces belong. If the system is unable to estimate the age group, the highest brightness level is used.

To report the gender of the detected face whether it is male or female the left-most or right-most phosphene in the first row is turned on respectively. These phosphenes are used to show the gender of unknown people. If more than one unknown face is detected in the frame, the gender of the person closest to the middle of the frame is shown. The remaining phosphenes in the first row are assigned to known people. Each phosphene is used for each known person. The maximum number of identified people that the system can report is 13.

Five phosphenes in the bottom of the first column in the SPV are used to express the result of facial emotion recognition. These phosphenes are used to show one of the following emotions: neutral, happy, sad, angry, and surprised. If no expression is detected, none of these phosphenes is turned on. If more than one face is detected in the frame, the emotion of the person closest to the middle of the frame is shown.

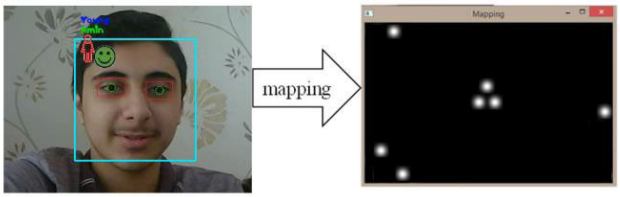

FIGURE 7. A practical implementation of the mapping.

The five phosphenes at the bottom of the last column are used to show the estimated distance between the detected person and the visually impaired person. If the size of the square frame around the face detected is smaller than a certain limit, the system considers that the detected person is far away and the phosphene at the lowest position is turned on. As a detected person gets closer to the user, one of the above phosphenes up to the fifth one is turned on. If more than one face is detected, the estimated distance for the person closest the center of the frame is shown. Finally, the 13 phosphenes in the middle of last row determine the number of known or

unknown people in the frame. An example of typical output is shown in Fig. 7.

V. EXPERIMENTAL RESULTS AND DISCUSSION

Many improvements for prosthetic vision proposed in recent years. They proposed different features for visually impaired people and used different resolutions of SPVs. Table 1 reports these works and compares the proposed system with them. This table reports the authors, date, the size of SPV, the number of graylevel and the feature proposed in each system. Some of these works did not proposed a new SPV, nevertheless they tested their improvement on VR head-mounted display. In one real case, the proposed system was tested with four Argus II users [27]. As shown in this table most of previous research are based on scene understanding, object and obstacles detection which improve visually impaired people mobility. There are very few works on human identification and recognition.

The proposed system conveys important information about the people around a visually impaired person and to make that person more comfortable to communicate with other people.

One important issue in comparing different methods is the resolution of SPVs. It is clear that phosphene images with higher resolutions are able to convey more information to the user, however, these phosphene images are not implementable in current technology of visual implants. The proposed system uses a number of phosphenes relatively close to the current commercialized prosthesis, which is not the case with some systems. Thus, the system that we propose can be considered for some next generation visual implant systems.

The important information that the proposed system transfers to the user are not found in the current prostheses and also, they are not provided by any other methods reported in Table 1. The proposed system tries to transform this information to the visually impaired person using schematic and symbolic patterns.

Another important issue is that the video processing unit of systems such as Argus II and other prosthetic vision can be managed to work in different modes of operation. Here we propose a mode of operation in which the visually impaired person copes with people at the office or at home and the mobility of that person in these environments is a lesser challenge because the user is already familiar with it.

Different functionalities of the proposed system are evaluated individually. Since the accuracy of the system is affected by its training data set, appropriate databases are required to assess each processing step. The proposed system is implemented using C++ and OpenCV libraries. Reported experimental results were obtained using a moderate performance computer equipped with a second generation core™i7-2640M processor and 8 GB of main memory. The proposed visual aid software chain running in this hardware offers a performance of 6 frames per second which is the same as the speed of Argus II [4]. A low-cost and energy efficient implementation of the proposed method was ported to a Raspberry Pi 4 B using OpenCV library run with the frame rate of 4.5 frames/second shows the feasibility of the hardware implementation of the proposed system.

TABLE 2. Accuracy of different gender detection methods.

Algorithms	Quantity
Eigenfaces [33]	95%
Fisherfaces [32]	87.5%
LBP [34]	80%

The experimental results for each phase of the proposed visual aid software are reported in the following.

A. FACE RECOGNITION ALGORITHM

The ORL database [45] containing 10 photos of 40 different people is used in the face recognition phase. After training the face recognition algorithm, a label is assigned to each identified person in the database. This label is returned as the output of the algorithm, and if a face is not identified, a negative label is returned. Fisherfaces algorithm is applied on the preprocessed images where 80% of the images were used for training and 20% for testing. The accuracy of this method is 90%.

B. GENDER DETECTION ALGORITHM

The accuracy of the gender detection algorithm is evaluated using 200 images from the FaceScrab database. These images comprise 100 male and 100 female samples that were randomly selected. Training was performed with a random selection of 80% of the data set whereas the remaining 20 percent is used for testing. The results of three different gender recognition algorithms, namely, LBP [34], Fisherfaces [32], and eigenfaces [33] on the FaceScrab database [39] are shown in Table 2. According to these results, the accuracy of the eigenfaces algorithm is far higher than the other algorithms.

C. AGE ESTIMATION ALGORITHM

To evaluate the performance of the proposed age estimation algorithm, the FG-Net [46] database was used. FG-Net comprises 1002 images of 82 people, with 6 to 12 images for each person at different ages from 0 to 69 years and each photo is labeled with the chronological age. 802 images in the database were used for training and 200 for testing. The average difference between the estimated and actual ages for each person in the data collection was calculated as the output error. Table 3 compares the mean absolute error (MAE) of the proposed method with two other well-known methods as well as two instances of human estimation tests. As reported in this table, the proposed algorithm outperforms the others as it offers the lowest MAE.

Table 4 reports the percentages of different amounts of differences between the chronical and predicted ages in the proposed method. As shown in this table, 30% of the

TABLE 3. MAE comparison of different age estimator algorithms on the FG-NET database.

Algorithms	MAE
Proposed method with 60 neurons in hidden layer	3.54
IIS-LLD(Gaussian)	5.77
CPP with 400 neurons in hidden layer	4.76
Human test 1	8.13
Human test 2	6.23

TABLE 4. The accuracy of age estimation algorithm based on proximity to the chronological age.

Distance between predicted age and chronological age (year)	Number of samples (precision)
0	30%
1	18.5%
2	13%
3	9.5%
4	8.5%
5	3.5%
6	3.5%
7	4%
8	3%
9	1.5%
10	1%
Over 10 years	4%

estimates were exactly the same as the true chronological ages and 18.5% fell within a one-year deviation.

D. FACIAL EXPRESSION RECOGNITION ALGORITHM

In the facial emotion recognition, testing and training were conducted on the Cohn-Kanade database, which is one of the most comprehensive databases used in facial expression research [47]. This database includes images of 100 students aged between 18 and 30 years. For each person, there is a sequence of black and white images with 480 × 640 resolution.

Five natural facial expressions, namely, natural, happy, sad, surprised, and angry, were selected for training the emotion recognition algorithm. A set of 347 samples in five classes were used; and 20% of the samples were used for testing and selecting the best algorithm. emotion recognition algorithm.

As mentioned before, different facial expression methods were tested to select an appropriate method based on computational complexity, accuracy and code availability. Table 5 shows these results. The original LBP feature extraction method and the LDA classification method were selected due to their low computational complexity even if they are not the most accurate.

E. GENERAL TESTING OF THE SYSTEM

A database of 500 images containing 250 male and female actors randomly selected from the FaceScrab database to evaluate the whole system. These images are not used in the training phase.

TABLE 5. The implementation results of accuracy of different facial expression recognition algorithms.

Algorithms	Accuracy
AAM + SVM	92%
LDA + Original LBP [40]	84.18%
LDA + Variance-based LBP [34]	83%
LDA + Circular LBP [48]	82%
Fisherfaces	60%
LBPfaces	59%
Eigenfaces	54.18%

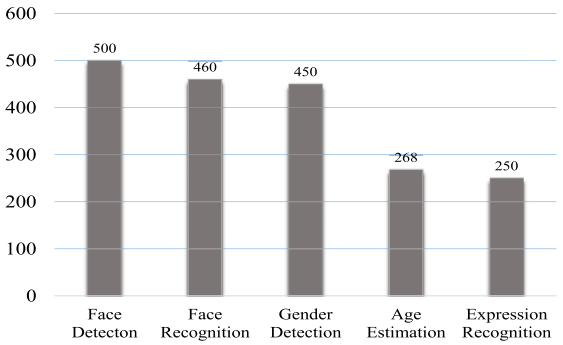

FIGURE 8. Accuracy graph of each algorithm for 500 sample images that are not part of the training set.

The accuracy of the face detection algorithm for images never seen by the system was 100%. Among the 500 images, 30 were from known people in the training phase; 27 of which were recognized successfully. On the other hand, among the other 470 unknown images, 37 were identified incorrectly.

For gender detection, 92.8% of the men's and 87.2% of the women's images were successfully classified.

In the age estimation phase, 500 images of celebrities were used and 53.6% of them were properly classified in one of three reported age groups (children, young and old). The error was mostly observed in the adult age groups due to deliberate changes in the faces such as make-up, cosmetic operation, etc.

In the facial expression recognition phase, since the system was trained with images under controlled lighting conditions, the accuracy was reduced. The system could correctly recognize emotions in 250 cases from a set of 500 samples.

The overall results for different phases in the proposed system are reported in Fig. 8.

VI. CONCLUSION

The present study aims at helping developers of smart data processing systems for visual implants that could be offered to visually impaired people. Due to the limited number of microelectrodes in visual stimulators, different machine vision and image processing algorithms were used to improve the artificial vision in the proposed system. The proposed system implements face detection and identification, gender detection, age estimation, and facial expression.

The proposed system provides more optional features to current visual prosthesis. The software of the proposed system can be integrated in the video processing unit of current prosthesis system such as Argus II. It can be programmed to switch between the previous and new mode of operation.

With advances in computer vision processing and feedback from visually impaired users, newer versions of the software can be provided without changes to the hardware. More information is shown in the phosphene images, and also, more accurate algorithms will be used in the improved versions.

The proposed system can be used for different kinds of vision prosthesis including retina and visual cortex implants. A visually impaired person using the proposed system must be trained to understand the information shown by each group of phosphenes.

Clearly, if greater number of electrodes becomes available in visual implants, the number of phosphenes would increase, data could be transferred with better resolution and accuracy, and more sophisticated vision modes could be offered.

ACKNOWLEDGMENT

The authors would like thank Mr. Julien Posso for assistance with implementing the code of the proposed system in Raspberry Pi B.

REFERENCES

[1] *Visual Impairment and Blindness*, Fact Sheet, World Health Organization, Geneva, Switzerland, 2014, pp. 2009–2014.
[2] N. Han, S. Srivastava, A. Xu, D. Klein, and M. Beyeler, "Deep learning–based scene simplification for bionic vision," in *Proc. Augmented Hum. Conf.*, Feb. 2021, pp. 45–54.
[3] J. de Ruyter van Steveninck, T. van Gestel, P. Koenders, G. van der Ham, F. Vereecken, U. Güçlü, M. van Gerven, Y. Güçlütürk, and R. van Wezel, "Real-world indoor mobility with simulated prosthetic vision: The benefits and feasibility of contour-based scene simplification at different phosphene resolutions," *J. Vis.*, vol. 22, no. 2, p. 1, Feb. 2022.
[4] D. Avraham, J.-H. Jung, Y. Yitzhaky, and E. Peli, "Retinal prosthetic vision simulation: Temporal aspects," *J. Neural Eng.*, vol. 18, no. 4, Aug. 2021, Art. no. 0460d9.
[5] X. Xia, X. He, L. Feng, X. Pan, N. Li, J. Zhang, X. Pang, F. Yu, and N. Ding, "Semantic translation of face image with limited pixels for simulated prosthetic vision," *Inf. Sci.*, vol. 609, pp. 507–532, Sep. 2022.
[6] D. Avraham and Y. Yitzhaky, "Effects of depth-based object isolation in simulated retinal prosthetic vision," *Symmetry*, vol. 13, no. 10, p. 1763, Sep. 2021.
[7] M. Beyeler and M. S. Garcia, "Towards a smart bionic eye: AI-powered artificial vision for the treatment of incurable blindness," *J. Neural Eng.*, vol. 19, no. 6, p. 063001, 2022.
[8] M. L. Italiano, T. Guo, N. H. Lovell, and D. Tsai, "Improving the spatial resolution of artificial vision using midget retinal ganglion cell populations modeled at the human fovea," *J. Neural Eng.*, vol. 19, no. 3, Jun. 2022, Art. no. 035002.
[9] W.-J. Chang, L.-B. Chen, C.-H. Hsu, J.-H. Chen, T.-C. Yang, and C.-P. Lin, "MedGlasses: A wearable smart-glasses-based drug pill recognition system using deep learning for visually impaired chronic patients," *IEEE Access*, vol. 8, pp. 17013–17024, 2020.
[10] Y. Ito, C. Premachandra, S. Sumathipala, H. W. H. Premachandra, and B. S. Sudantha, "Tactile paving detection by dynamic thresholding based on HSV space analysis for developing a walking support system," *IEEE Access*, vol. 9, pp. 20358–20367, 2021.
[11] S. Niketeghad and N. Pouratian, "Brain machine interfaces for vision restoration: The current state of cortical visual prosthetics," *Neurotherapeutics*, vol. 16, no. 1, pp. 134–143, Jan. 2019.
[12] P. M. Lewis, L. N. Ayton, R. H. Guymer, A. J. Lowery, P. J. Blamey, P. J. Allen, C. D. Luu, and J. V. Rosenfeld, "Advances in implantable bionic devices for blindness: A review," *ANZ J. Surg.*, vol. 86, no. 9, pp. 654–659, Sep. 2016.
[13] S. Ha, M. L. Khraiche, A. Akinin, Y. Jing, S. Damle, Y. Kuang, S. Bauchner, Y.-H. Lo, W. R. Freeman, G. A. Silva, and G. Cauwenberghs, "Towards high-resolution retinal prostheses with direct optical addressing and inductive telemetry," *J. Neural Eng.*, vol. 13, no. 5, Aug. 2016, Art. no. 056008.
[14] P. M. Lewis, H. M. Ackland, A. J. Lowery, and J. V. Rosenfeld, "Restoration of vision in blind individuals using bionic devices: A review with a focus on cortical visual prostheses," *Brain Res.*, vol. 1595, pp. 51–73, Jan. 2015.
[15] J. D. Weiland and M. S. Humayun, "Retinal prosthesis," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 5, pp. 1412–1424, May 2014.
[16] E. Bloch, Y. Luo, and L. da Cruz, "Advances in retinal prosthesis systems," *Therapeutic Adv. Ophthalmol.*, vol. 11, Jan. 2019, Art. no. 251584141881750.
[17] J. Jang, H. Kim, Y. M. Song, and J. U. Park, "Implantation of electronic visual prosthesis for blindness restoration," *Opt. Mater. Exp.*, vol. 9, no. 10, pp. 3878–3894, Oct. 2019.
[18] *Second Sight Company*. Accessed: Dec. 13, 2020. [Online]. Available: http://secondsight.com
[19] *Vivani Company*. Accessed: Jul. 28, 2023. [Online]. Available: http://vivani.com
[20] D. D. Zhou, J. D. Dorn, and R. J. Greenberg, "The Argus II retinal prosthesis system: An overview," in *Proc. IEEE Int. Conf. Multimedia Expo. Workshops (ICMEW)*, Jul. 2013, pp. 1–6.
[21] K. Stingl, K. U. Bartz-Schmidt, D. Besch, A. Braun, A. Bruckmann, F. Gekeler, U. Greppmaier, S. Hipp, G. Hörtdörfer, C. Kernstock, A. Koitschev, A. Kusnyerik, H. Sachs, A. Schatz, K. T. Stingl, T. Peters, B. Wilhelm, and E. Zrenner, "Artificial vision with wirelessly powered subretinal electronic implant alpha-IMS," *Proc. Roy. Soc. B, Biol. Sci.*, vol. 280, no. 1757, Apr. 2013, Art. no. 20130077.
[22] T. L. Edwards, "Assessment of the electronic retinal implant alpha AMS in restoring vision to blind patients with end-stage retinitis pigmentosa," *Ophthalmology*, vol. 125, no. 3, pp. 432–443, Mar. 2018.
[23] G. S. Brindley and W. S. Lewin, "The sensations produced by electrical stimulation of the visual cortex," *J. Physiol.*, vol. 196, no. 2, pp. 479–493, May 1968.
[24] M. Sanchez-Garcia, R. Martinez-Cantin, and J. J. Guerrero, "Semantic and structural image segmentation for prosthetic vision," *PLoS ONE*, vol. 15, no. 1, Jan. 2020, Art. no. e0227677.
[25] A. Lozano, J. S. Suárez, C. Soto-Sánchez, J. Garrigós, J. J. Martínez-Alvarez, J. M. Ferrández, and E. Fernández, "Neurolight: A deep learning neural interface for cortical visual prostheses," *Int. J. Neural Syst.*, vol. 30, no. 9, Sep. 2020, Art. no. 2050045.
[26] H. Jiang, H. Li, J. Liang, and X. Chai, "A hierarchical image processing strategy for artificial retinal prostheses," in *Proc. Int. Conf. Artif. Intell. Comput. Eng. (ICAICE)*, Oct. 2020, pp. 359–362.
[27] A. Kartha, R. Sadeghi, M. P. Barry, C. Bradley, P. Gibson, A. Caspi, A. Roy, and G. Dagnelie, "Prosthetic visual performance using a disparity-based distance-filtering system," *Transl. Vis. Sci. Technol.*, vol. 9, no. 12, p. 27, Nov. 2020.
[28] F. Guo, Y. Yang, and Y. Gao, "Optimization of visual information presentation for visual prosthesis," *Int. J. Biomed. Imag.*, vol. 2018, pp. 1–12, Mar. 2018.
[29] N. Parikh, L. Itti, M. Humayun, and J. Weiland, "Performance of visually guided tasks using simulated prosthetic vision and saliency-based cues," *J. Neural Eng.*, vol. 10, no. 2, Feb. 2013, Art. no. 026017.
[30] H. M. Mohammadi, E. Ghafar-Zadeh, and M. Sawan, "An image processing approach for blind mobility facilitated through visual intracortical stimulation," *Artif. Organs*, vol. 36, no. 7, pp. 616–628, Jul. 2012.
[31] P. Viola and M. J. Jones, "Robust real-time face detection," *Int. J. Comput. Vis.*, vol. 57, no. 2, pp. 137–154, May 2004.
[32] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, Jul. 1997.
[33] M. Turk and A. Pentland, "Eigenfaces for recognition," *J. Cognit. Neurosci.*, vol. 3, no. 1, pp. 71–86, Jan. 1991.
[34] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Aug. 2002.
[35] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. CVPR*, Dec. 2001, p. 1.
[36] V. Bruce, A. M. Burton, E. Hanna, P. Healey, O. Mason, A. Coombes, R. Fright, and A. Linney, "Sex discrimination: How do we tell the difference between male and female faces?" *Perception*, vol. 22, no. 2, pp. 131–152, Feb. 1993.

[37] X. Geng, C. Yin, and Z.-H. Zhou, "Facial age estimation by learning from label distributions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 10, pp. 2401–2412, Oct. 2013.
[38] M. F. Valstar, I. Patras, and M. Pantic, "Facial action unit detection using probabilistic actively learned support vector machines on tracked facial point data," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR) Workshops*, Sep. 2005, p. 76.
[39] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 2037–2041, Oct. 2006.
[40] T. Ojala, M. Pietikäinen, and T. Mäenpää, "A generalized local binary pattern operator for multiresolution gray scale and rotation invariant texture classification," in *Proc. Int. Conf. Adv. Pattern Recognit.* Berlin, Germany: Springer, 2001, pp. 399–408.
[41] J. Kumari, R. Rajesh, and K. M. Pooja, "Facial expression recognition: A survey," *Proc. Comput. Sci.*, vol. 58, pp. 486–491, Jan. 2015.
[42] S. C. Chen, G. J. Suaning, J. W. Morley, and N. H. Lovell, "Simulating prosthetic vision: I. Visual models of phosphenes," *Vis. Res.*, vol. 49, no. 12, pp. 1493–1506, Jun. 2009.
[43] J. Wang, X. Wu, Y. Lu, H. Wu, H. Kan, and X. Chai, "Face recognition in simulated prosthetic vision: Face detection-based image processing strategies," *J. Neural Eng.*, vol. 11, no. 4, Jun. 2014, Art. no. 046009.
[44] W. L. D. Lui, D. Browne, L. Kleeman, T. Drummond, and W. H. Li, "Transformative reality: Improving bionic vision with robotic sensing," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2012, pp. 304–307.
[45] F. S. Samaria and A. C. Harter, "Parameterisation of a stochastic model for human face identification," in *Proc. IEEE Workshop Appl. Comput. Vis.*, Dec. 1994, pp. 138–142.
[46] A. Lanitis, "Comparative evaluation of automatic age progression methodologies," *EURASIP J. Adv. Signal Process.*, vol. 2008, no. 1, 2008, Art. no. 239480.
[47] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The extended cohn-kanade dataset (CK+): A complete dataset for action unit and emotion-specified expression," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops*, Jun. 2010, pp. 94–101.
[48] L. Liu, P. Fieguth, G. Zhao, M. Pietikäinen, and D. Hu, "Extended local binary patterns for face recognition," *Inf. Sci.*, vols. 358–359, pp. 56–72, Sep. 2016.

HOSSEIN MAHVASH MOHAMMADI received the B.S. and M.S. degrees in computer engineering from Amir Kabir University, Tehran, Iran, in 1992 and 1997, respectively, and the Ph.D. degree in electrical engineering from École Polytechnique de Montréal, QC, Canada, in 2009. He was a Postdoctoral Fellow of the Polystim Neurotechnologies Laboratory, Polytechnique Montréal, in 2010, and a Postdoctoral Fellow of the Imaging and Orthopaedics Research Laboratory, Ecole de Technologie Superieure, Montréal, in 2011. He was a Visiting Researcher with the Center of Intelligent Machines, McGill University, in 2004, and a Visiting Researcher with École Polytechnique de Montréal, in 2020, and a Visiting Professor with École Polytechnique de Montréal, in 2022. Since 2012, he has been with the University of Isfahan, where he is currently an Assistant Professor with the Department of Computer Engineering. His research interests include image and video processing, computer vision, computer graphics, 3D rendering, visual implants, medical imaging, biomedical engineering, and deinterlacing.

MOHAMMAD HADI EDRISI, photograph and biography not available at the time of publication.

YVON SAVARIA (Fellow, IEEE) received the B.Ing. and M.Sc.A. degrees in electrical engineering from Polytechnique Montréal, in 1980 and 1982, respectively, and the Ph.D. degree in electrical engineering from McGill University, in 1985. He was a Consultant or was sponsored for carrying research by Bombardier, CNRC, DesignWorkshop, DREO, Ericsson, Genesis, Gennum, Huawei, Hyperchip, Intel, ISR, Kaloom, LTRIM, Medvalgo, Miranda, MiroTech, Nortel, Octasic, PMC-Sierra, Technocap, Thales, Tundra, and Wavelitepatents. He was the Thesis Advisor of 160 graduate students who completed their studies. Since 1985, he has been with Polytechnique Montréal, where he is currently a Professor with the Department of Electrical Engineering. He is affiliated with the Hangzhou Innovation Institute, Beihang University. He holds 16 patents and has published 170 journal articles and 450 conference papers. He is currently involved in several projects that relate to aircraft embedded systems, radiation effects on electronics, asynchronous circuits design and test, green IT, wireless sensor networks, virtual networks, software-defined networks, machine learning, computational efficiency, and application-specific architecture design. His research interests include microelectronic circuits and microsystems, such as testing, verification, validation, clocking methods, defect and fault tolerance, the effects of radiation on electronics, high-speed interconnects and circuit design techniques, CAD methods, reconfigurable computing and the applications of microelectronics to telecommunications, aerospace, image processing, video processing, radar signal processing, and digital signal processing acceleration. He is a member of Regroupement Stratégique en Microélectronique du Québec (RESMIQ), Ordre des Ingénieurs du Québec (OIQ), and the CMC Microsystems Board. In 2001, he was awarded the Tier 1 Canada Research Chair on design and the architectures of advanced microelectronic systems that he held until June 2015. He was a recipient of the Synergy Award of the Natural Sciences and Engineering Research Council of Canada. He was the Program Co-Chairman of NEWCAS'2018.