We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

Interested in publishing with us?
Contact book.department@intechopen.com

Imaging Systems in Assistive Technology

Miloš Klíma and Stanislav Vítek
Czech Technical University in Prague
Czech Republic

1. Introduction

The working age population, which is conventionally defined as aged between 15 and 64 years, would start to decline as of 2010 and, over the period of 50 years, it would drop by 15% in the EU. In 2060, there would be more than twice as many elderly than children. In 2008, there were about three and a half times as many children as very old people (above 80). In 2060, children would still outnumber very old persons, but by a small margin: the number of very old people would amount to 80% of the number of children.

These numbers are very significant – more money will be spent on care for the sick and elderly people. And what is far more important, there would be not enough labor force, which should take care of these people. In this situation can help the assistive technologies, covering a broad range of different techniques and approaches streaming to the same aim – support and assistance for elderly and handicapped persons. It is a typical interdisciplinary field highly demanding not only towards designers and manufacturers but also towards users.

The labor shortage can be largely solved by applied multimedia imaging systems. Remote monitoring can reduce the amount of recurring admissions to hospital, faciliate more efficient clinical visits with objective results, and may reduce the length of a hospital stay forindividuals who are living at home. Telemonitoring can also be applied on long-term basis to elderly persons to detect gradual deterioration in their health status, which may imply a reduction in their ability to live independently. In this chapter follows general overview and possible applications of such a sytems and particularly inteligent CCTV (Closed Circuit TV) surveillance systems for monitoring of handicaped people, and an aids for everyday use, utilizing multimedia means in any form.

2. Imaging system in general

The imaging system in general is demonstrated in the block diagram on Fig. 1. The first part of system is formed by an optical system. The optical system transforms 3D information from the scene into 2D planar projection on an image sensor (2D distribution of irradiation or illumination). In some cases there are some additional elements inserted between the optical system and sensor – such as image amplifier, shutter, various filters etc. The image sensor performs a conversion of optical projected image into its electrical representation – electro-optic conversion. The image sensor is a 2D array of pixels (picture elements) and supply electrical signal to one or several outputs. The video electrical signal passes through the preprocessing and source coding blocks. Final video signal can be transmitted through

the communication channel or retrieved. The separate block is an image processing block providing the selected processing tools – motion detection, object detection and identification, object parameter measurement etc. The final block is an image display converting the electrical representation into the 2D or stereoscopic optical image viewed by an observer.
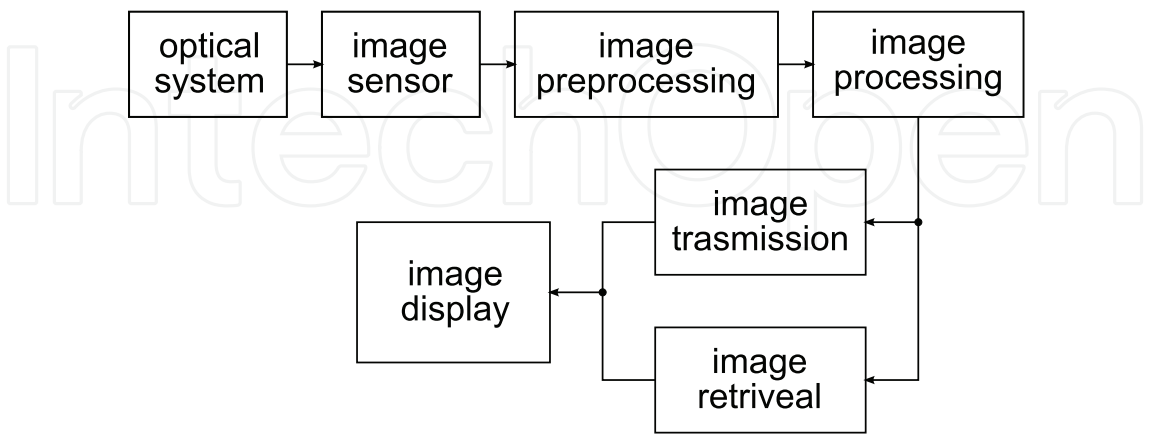
Fig. 1. Imaging system in general

3. Optical system and additional elements

The optical system is formed by an objective – usually as a combination of lenses (refractive optics) but there are some other configurations available – reflective (mirror) optics and diffraction elements. The selection is based upon selected spectral range and particular purpose. Relevant spectral transparency is limiting for lenses, spectral reflexivity for reflective optics. There is a large variety of available materials [X1] for refractive optics e.g. glass for VIS, ZnSe and Ge for IR, quartz and sapphire for UV. In the X-ray range the optical imaging system is constructed as mirror optics. In special terrestrial applications we have to take into account also the transparency of the atmosphere e.g. in X-ray imaging.

The important parameters of optical system are viewing angle, resolution and aberrations. The resolution is frequently described in a number of lines per mm but more precise definition is done by the MTF (Modulation Transfer Function) or PSF (Point Spread Function). The MTF is a module of optical transfer function in spatial domain; PSF is a spatial impulse response of the system. Apart of that the geometrical image distortions (barrel or cushion type) and vignetting are affecting the final image quality. As an integral part of optical system is an aperture. It is characterized by F-number. The function of aperture is multiple. At first it controls the level of sensor illumination (irradiation). The other aperture function is to set the DOF (Depth of Field, Depth of Focus) and the last one affects the resolution (large aperture – larger impact of abberations, small aperture – larger impact of diffraction).

The first additional element is an image amplifier with optical image input and optical image output. There have been used vacuum versions for long time but now the MCP (Microchannel Plate) is widely applied – see Fig. 2.

The other additional elements in the optical system are absorption or interference filters modifying the spectral transparency, polarization filters, etc.
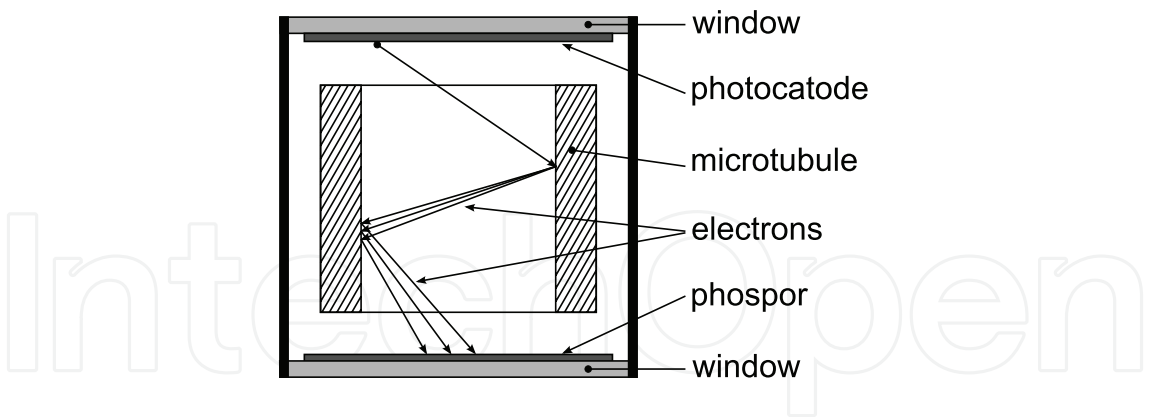
Fig. 2. Microchannel plate.

4. Image sensing

The image sensor is a fundamental part of imaging system. The sensor itself is always a planar array of sensing elements and it means that the sensor performs a 2D spatial sampling of projected image. The 2D sampling can affect the sampled image quality in the same way as it is in 1D sampling in signal domain. The resolution (or MTF and PSF) of optical imaging system should be matched to the spatial sampling raster of sensor otherwise some additional image distortions can be introduced. The spatial sampling frequency value has to fulfill the Shannon sampling theorem (sampling frequency must be at least twice higher than the highest spatial frequency in the image). If not the aliasing effects will appear – very well-known effect called moire. The problem is even more difficult in the color-sensing single-chip image sensors where the color splitting system is created by a CFA (Color Filter Array) in planar version. In this case the coinciding color rasters provides lower spatial sampling frequency than relates to the total pixel number. The example is shown on Fig. 3 – the most frequently used Bayer RGB CFA. Sensor manufactures eliminate this problem by the implementing of 2D anti-aliasing low-pass spatial filter front of the sensor (two optically birefringent plates) in order to limit a maximum spatial frequency in the projected image. For details see Lukac (2009).
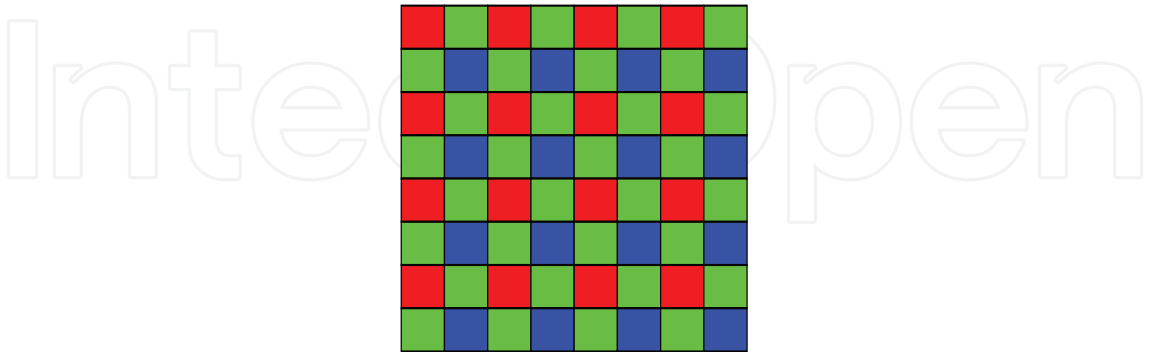

Fig. 3. RGB CFA - Bayer structure.

The CFA is realized as an array of thin-film interference filters and their spectral responses are not matched to a particular RGB color system and the RGB signals should be transformed to some commonly used.

Recently there are available two basic types of sensors – CMOS (Complementary Metal-Oxide-Semiconductor) called after the standard memory chip technology and CCD (Charge Transfer Device). The CMOS sensing elements are a PN or PIN photodiodes usually scanned in the XY commuting array structure – see Fig. 4.
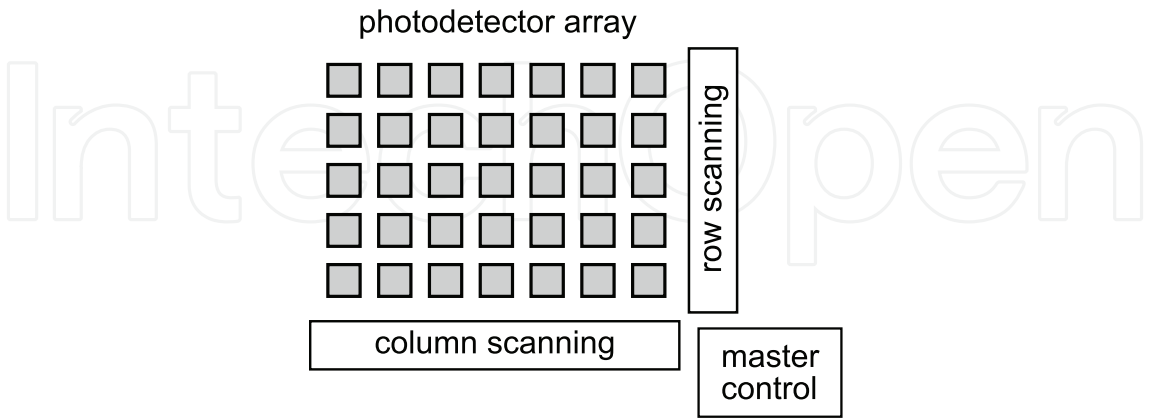
photodetector array

Fig. 4. XY CMOS sensor structure and basic commuting principle.

The CMOS sensors provide high flexibility of scanning sequence (random access, scanning of sensor part) and driving and signal processing circuitry is included in the sensor chip. Moreover the CMOS sensors can offer the function called binning – parallel connection of neighboring pixels in order to increase sensor sensitivity. In order to avoid the problem with the planar CFA 2D sampling the FOVEON® introduced a different vertical color splitting structure where the layers of silicon itself are used as color filters – see Fig. 5.
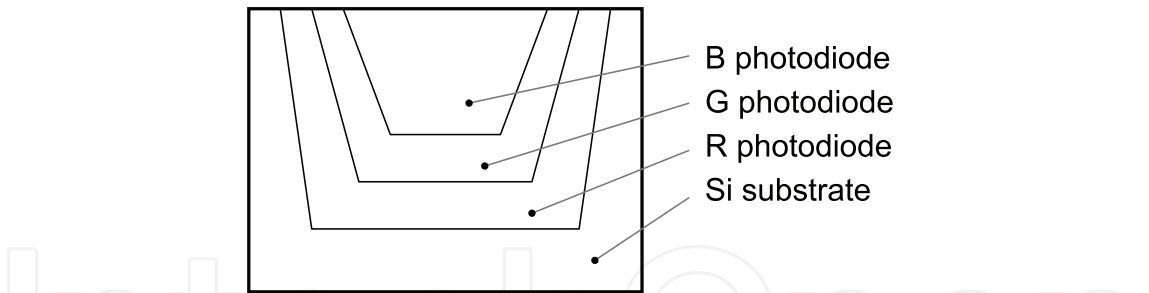

Fig. 5. FOVEON® CFA structure.

CCD sensors are based upon a different principle. The pixel is formed by a MOS capacitor and the signal quantity is a size of charge packet accumulated in this capacitor. The charge packet size is proportional to local pixel illumination. The pixels are organized in rows forming a charge packet accumulating and transferring structure – see Fig. 6. The vertically oriented sensing and transferring lines creates a 2D pixel array. In the FT (Frame Transfer) configuration the 2D charge-packet distribution (frame) is transferred into the similar 2D pixel array. This part is not photosensitive and is used as a memory with joint line shift register array. When the image has been transferred to the memory part the stored image frame is read out line-by-line and pixel-by-pixel. In such a way the image scanning procedure is fixed and cannot be changed and pixels must be read out in given order. Recently the image sensors are manufactured in resolutions from millions to tens of millions pixels.
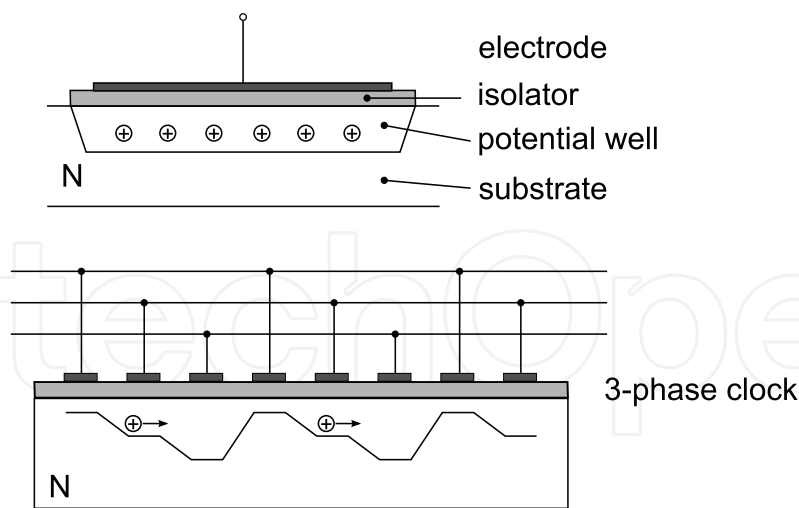
Fig. 6. CCD structure.

Specialized sensors are used esp. in IR spectral regions such as pyroelectric detector array, microbolometric array, microthermocouple array, photoresistor array and also CCD sensors on e.g. HgCdTe semiconductor (different from silicon). Another approach is to convert invisible optical radiation outside of VIS range into VIS applying a proper phosphor or scintillator. It mostly applied for high-energy photons in UV, X-ray a gama spectral range.

5. Image reproduction

Recently image displays are based upon several physical principles. Recent generation displays are LCD (Liquid Crystal Display), plasma, DLP (DMD) and OLED (Organic Light Emitting Diode). The coming generation seems to be FED (Field Emission Display) and SED (Surface Emission Display). Printers dominate the ink-jet and laser types.

The color displays create a color image in similar way as the image sensors split a color image into RGB trichromatic representation. In principle, no trichromatic display with three basic colors can reproduce all existing colors esp. highly saturated. Fig. 7 shows the examples of existing structures. White color is used in some of them to avoid requirement of maximum signals for all three RGB channels and to guarantee correct reproduction of white.
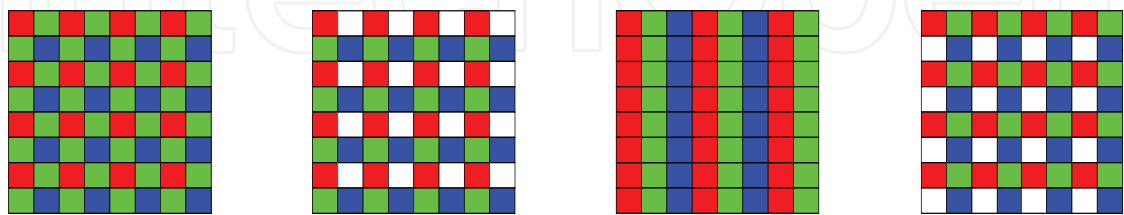

Fig. 7. Structure of reproduction rasters in trichromatic displays.

The LCD pixel functions as an electrically-controlled rotator of light polarization plane. In the absence of external electric field the elongated rod-type molecules of liquid crystal create a twisted structure performing the polarization plane rotation. When the external electric

field (driving voltage) is applied the liquid crystal molecules orientation will follow the direction of electric field providing no polarization plane rotation. Combining this element with polarizer(s) provide a required light intensity control – see Fig. 8. The LCD panel can be designed as transparent or reflexive versions.

The LCD does not generate light but can only modulate light from an external light source (it functions as a SLM Spatial Light Modulator or valve). Therefore the LCD uses a luminescent panels or LED array in order to provide required light. LCDs are frequently applied in image projectors where the external light source is usually a xenon-filled discharge tube. The LCD pixel array is multiplexed in a standard XY (column and row) configuration. In TFT (Thin Film Transistor) active displays the LCD pixel array is covered an array of thin film transistors. Each LCD pixel is equipped with one or two TFTs creating a one-bit memory holding information about a required LCD pixel state. It stays until is changed. This technique improves highly LCD performance (brightness, contrast).
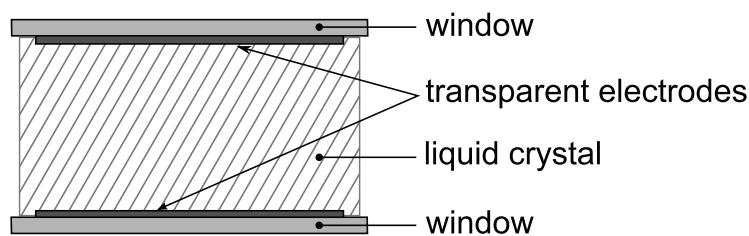

Fig. 8. LCD display principle.

Plasma panel is based upon a pixel array of micro discharge tubes. The discharge tube is filled with a mixture of rare gases e.g. Ne, Xe. The discharge emits UV radiation and excites a layer of phosphor. There are applied different phosphors for R, G and B pixels. Fig. 9 demonstrates the principle. The scanning structure is again XY array but no additional one-bit pixel memory is required. According to the discharge tube VA characteristic there are two stable states – on and off. So the one-bit memory is provided by the discharge itself.
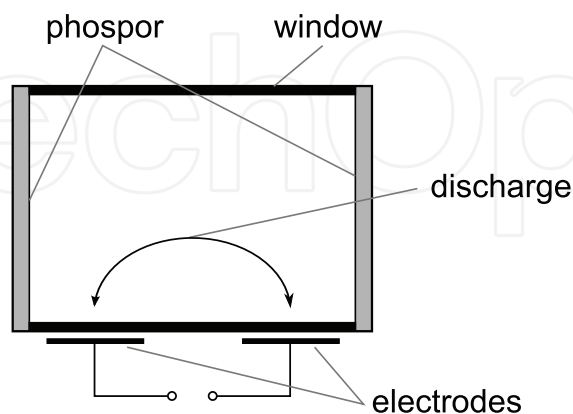

Fig. 9. Plasma display principle.

OLED panels are based upon organic semiconductor materials. In such a material the LED array is created including commuting structure. They are already widely applied in

mobile phones and other small screen applications. The advantage of OLED is very wide emitted light spectrum (reproduction of highly saturated colors), transparency of material and mechanical flexibility.

DLP (Digital Light Projector) is equipped with the display chip DMD (Digital Micromirror Device). It is an array of electrostatically deflected micromirrors. Each mirror is controlled separately and its position is stored in a one-bit memory underneath until is changed – for illustration see Fig. 10. The DLP technology is applicable for the image projection to the largest available screens (tens of meters). The DMD is fast sufficiently to enable the time multiplexing of colors instead of isochronal reproduction of colors in all other systems. The time multiplexing of colors was applied in the first-generation color TV cameras before and during World War II.
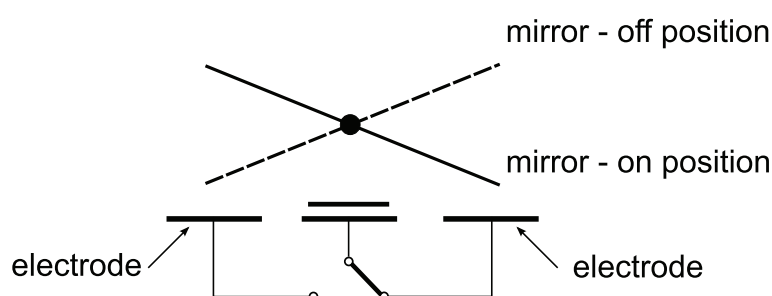

Fig. 10. DLP - DMD principle.

The plasma and DMD have only two states – on and off. Therefore for a gray-scale reproduction the PWM (Pulse Width Modulation) has to be applied for the light intensity control. Moreover using a sequence of impulses of proper length for each pixel one can realize a direct in-pixel D/A conversion.

Near-future displays seem to be FET and SED panels. In both cases it is in fact redesigned principle of old CRT (Cathode Ray Tube) with an electron beam modulation. In all three cases the electron beam is electrically modulated and after that creates an image in a phosphor layer. The source of electrons in CRTs is a heated cathode common for a whole CRT. In FEDs or SEDs there is a cold cathode generating electrons together with a modulation electrode in each pixel. Both principles are demonstrated on Fig. 11 and Fig. 12, respectively.

Stereoscopic (or 3D) image reproduction has been long-term exploited mostly in technical tasks (anaglyph, simulators etc.) and now is entering TV and games fields. The eye-multiplexing can be performed by color, polarization or switched multiplexing goggles/glasses or a display with a raster of cylindrical lenses.

6. Image preprocessing

The image preprocessing includes numerous functionalities to be performed on a raw image. We can list them in the following summary: correction of transfer functions (contrast enhancement) incl. histogram equalization, de-noising, sharpening (edge enhancement), white balance, color system transformation, de-vignetting, etc.
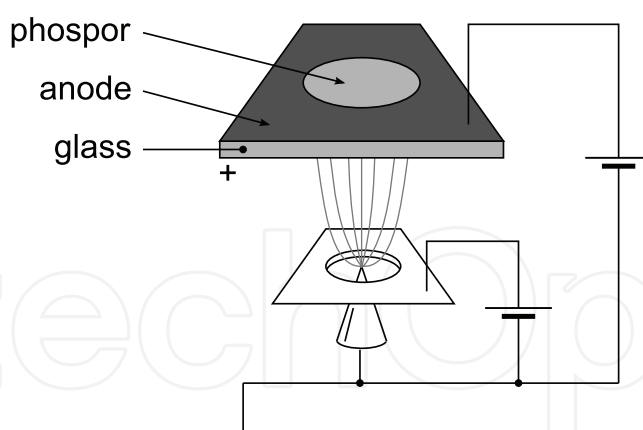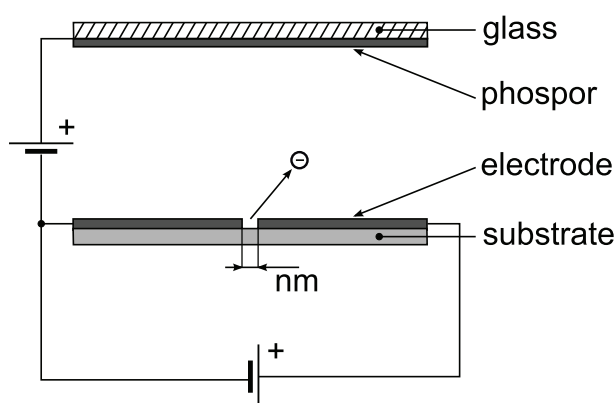
Fig. 11. FED display panel.

Fig. 12. SED principle.

7. Image/video compression and retriveal

The image compression or source coding is a vital task in imaging systems. The information content (size of file, required bitrate) is very high and it can be reduced significantly. Generally, using lossless compression techniques we can approach a limit by the redundancy. If the compression requirements are higher lossy compression techniques have to be applied. The values of bitrates can be demonstrated on the oldest digital TV standard ITU-R 601 defining SDTV (Standard Definition TV). It requires 216 Mbits/sec for 8bit quantization (720x576 pixels, 4:3, 25 frames/sec, signals Y-13.5 Msamples/sec, R-Y, B-Y - 6.5 Msamples/sec). Recently additional digitizing rasters have been defined - HDTV (High Definition TV, 1k, 16:9), Digital Cinema (2k, 4k) a Super HiVision (8k).

At the beginning of color TV in forties and fifties in the last century several analog source compression standards were developed. Based upon the physiology of HVS (Human Visual System) the basic color signals RGB (required bandwidth 3 x 6 MHz) have been transformed into a combination of the luminance signal Y and two chrominance signals e.g. R-Y, B-Y. The bandwidth requirements of R-Y and B-Y signals are reduced significantly (Y - 6 MHz, R-Y, B-Y - 1.6 MHz). These input signals are consequently coded in a complex color TV signal. Two chrominance signals are modulated by a QAM (Quadrature Modulation) on a color subcarrier and then frequency multiplexed with the luminance signal. The analog

coding standards are NTSC (National Television System Committee, USA and Japan) and PAL (Phase Alternating Line, Europe). The SECAM (Sequential a Memoire) originated in France and uses FM modulation of two color subcarriers (France, Eastern Europe). The simplicity of analog source codec schemes leads to continuing exploitation in cable TVs and especially in specialized CCTV systems for technical use but they are disappearing from broadcasting.

Recently, the vast number of digital image compression standards has been developed and is applied. There are still-picture compression standards (JPEG, JPEG2000, TIFF, GIF, PNG, etc.) and video compression standards (MPEG2, MPEG4/Part 2, MPEG4/Part 10/H.264/AVC, WMP, etc.). Most of them are using lossy compression techniques leading to some additional distortions and artifacts. Fig. 13 demonstrate typical example of distortions of JPEG standard.

(a) Original image   (b) Distorted image

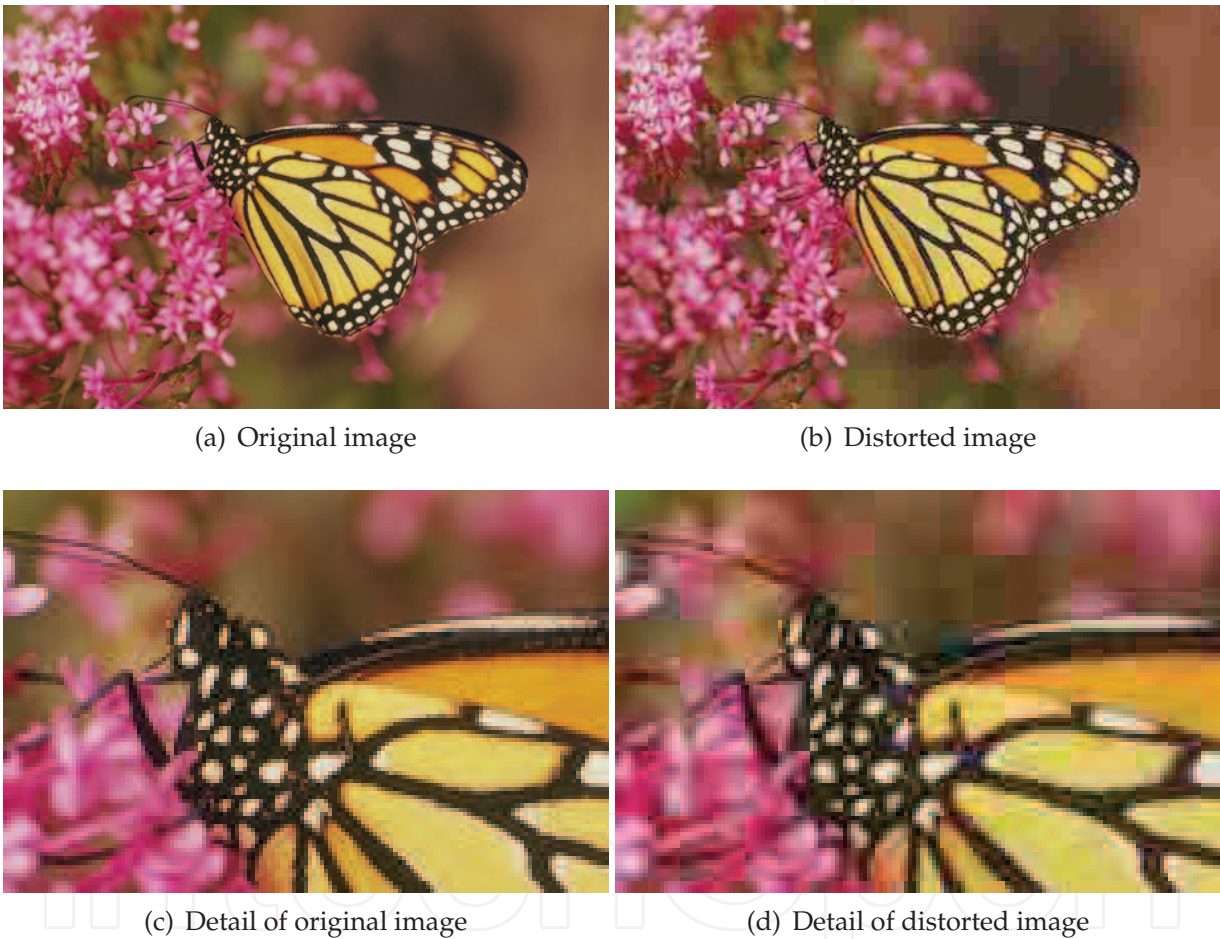(c) Detail of original image   (d) Detail of distorted image

Fig. 13. Example of distortions of JPEG standard.

The image compression standards are in fact procedures for applying image compression tools (techniques, methods). The first compression tool is given by a sampling raster. In color images the full resolution is provided only in the RGB representations. Otherwise the above mentioned signals Y, R-Y, B-Y are used. The full resolution is done in the Y channel only and the R-Y and B-Y signals are sampled in reduced sampling pattern (2 x lower in x direction, 2 x lower in y direction = 4 x in total). Among others there are two most important methods: DCT (Discrete Cosine Transform) with quantization and motion prediction. In all standards an image is split into blocks (JPEG, MPEG1,2 - 8 x 8 pixels, H.264/AVC variable

size) at the beginning. The DCT is applied on each block and it results into the set of spectral coefficients. The DCT concentrates energy (lowers redundancy) around the DC component and significantly reduces a number of coefficients to be transmitted. Each coefficient is consequently divided by relevant element of quantization matrix of same dimensions as the original block. The value of quantization matrix elements is perceptually optimized according to the subjective importance of each spectral coefficient. In such a way the number of coefficients is further reduced. Applying a zig-zag reading procedure for coefficients avoids a transmission of the most zero values. Therefore, the block information amount is reduced significantly and is content-dependent. Possible modification of quantization matrix gives a tool for image quality and required bitrate adjustment. In video the motion prediction technique is applied. The initial idea is to describe motion in image as a motion of blocks. The position of particular block is analyzed in the current and following blocks. The block position in another frame is searched through a search window and evaluated by selected block matching criterion. The block motion is than described by a motion vector. There are numerous other image compression techniques applied - VLC (Variable Length Codeword, e.g. Huffman coding), RLC (Run Length Coding), DPCM (Differential Pulse Code Modulation), WT (Wavelet Transform, JPEG 2000) etc.

8. Image/video transmissionn

As already mentioned there are many applications of imaging systems where the analog image compression and transmission is still applied. The analog video transmission is based on same principles as video broadcasting. The complex NTSC resp. PAL signal is VSB-AM (Vestigial Sideband – Amplitude Modulation) modulated on a particular RF carrier forming a TV channel. Because of historical reasons the structure of TV channels is different in different in different regions (USA and Japan, Western Europe, Central and Eastern Europe).

In eighties of the last century the digital TV started and there are two main standards – ATSC (USA) and DVB (Digital Video Broadcasting). They applied the MPEG 2 coding scheme and H.264/AVC coding scheme mostly for HDTV transmission. The ATSC uses the digital 8-state VSB-AM modulation. DVB has developed several broadcasting standards: DVB-T (terrestrial), DVB-S (satellite), DVB-C (cable), DVB-H (handheld), DVB-SH (satellite handheld) etc. Some of these standards are available in second generations: DVB-T2, DVB-S2. These standards differs in a channel coding scheme called FEC (Forward Error Correction) and type of RF carrier modulation according to the properties of particular transmission channel. Two examples have to be mentioned – DVB-T with OFDM (Orthogonal Frequency Division Multiplex) compensating an influence of multipath RF propagation and enabling a SFN (Single Frequency Network) transmitter operation. The other one is DVB-C accepting really multistate modulation up to 256 QAM. Another communication channel applied for image/video transmission is ADSL/VDSL approach trough a twisted-pair metal line. Internet has brought the IP TV streaming.

9. Distortions and artifacts, quality metrics

Image or video information is subjected to a broad variety of distortions during sensing, processing, compression, retrieving and transmission that image quality is a crucial question.

In assistive technologies we mostly expect an image assessment by a human observer but some approaches of machine vision can be implemented too.

There are two basic attitudes to the image quality assessment. The first one is purely oriented to the subjective perception and emotional impact of image - related to subjective quality. Generally, it is defined as QoE (Quality of Experience) against QoS (Quality of Service) usually defined purely technical. The second one can be called as a "feature-preserving" quality – related to detection, identification and classification of objects and events in image/video information. In both cases a lot of effort has been done in modeling of dependencies of subjective quality on some objective image parameters. The most reliable technique for the evaluation of image/video quality is a subjective assessment by a group of observers. There are numerous standards defined by ITU, ETSI, and SMTPE etc. describing practical procedures to be followed in subjective tests. On the other hand, there are numerous objective criteria expressing the image quality – from the simplest ones as MSE or PSNR up to much more sophisticated as SSIM (Structural Similarity Image Metric or VIF (Video Image Fidelity). These metrics can be classified as Full Reference, Reduced Reference and No Reference based upon the availability of original (reference) image.

10. Camera networks

Imaging system itself can only simply monitor the relatively limited space with no possibility to further processing. In order to be truly useful for assistive purposes, it's necessary bring imaging systems together into organized networks and systematically analyze if possible in real time video or image sequence and detect major events in terms of assistive applications.
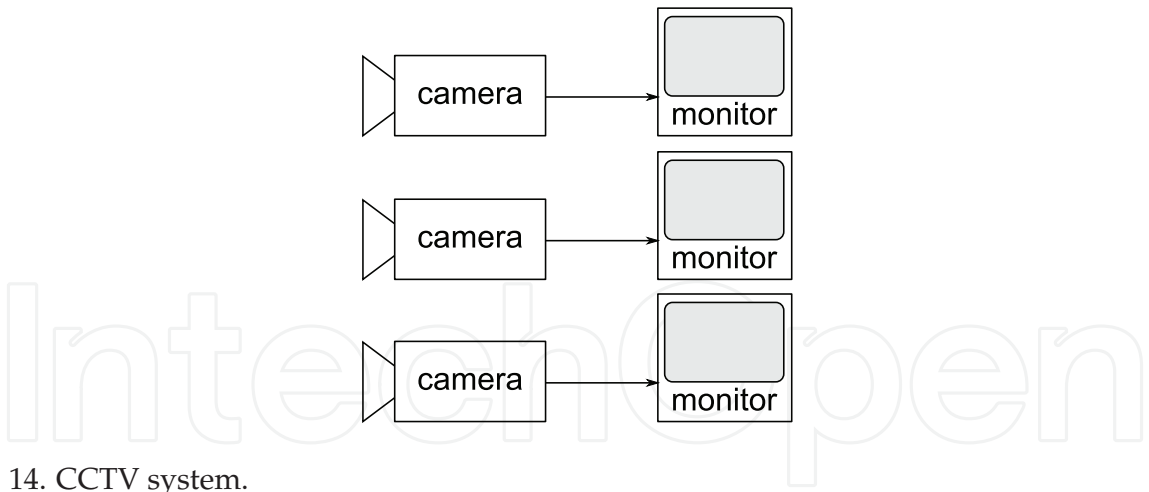

Fig. 14. CCTV system.

The simplest closed-circuit television (CCTV) system shown in Fig. 14 allowed an operator to observe roughly tens of different locations from many viewpoints at once. Systems can use both fixed on moving (PTZ, Pan-Tilt-Zoom) cameras – with PTZ cameras operator can change view of interest if necessary. More complex, centralized system where video from cameras is stored in at a central server which also distributes video for analysis, and to the user through a computer interface is shown in Fig. 15. The most sophisticated decentalized architecture is shown in Fig. 16. So called *smart cameras* are carried out video storage and data processing so bandwidth requirements are reduced and system with distributed metadata can be accessed from multiple locations.
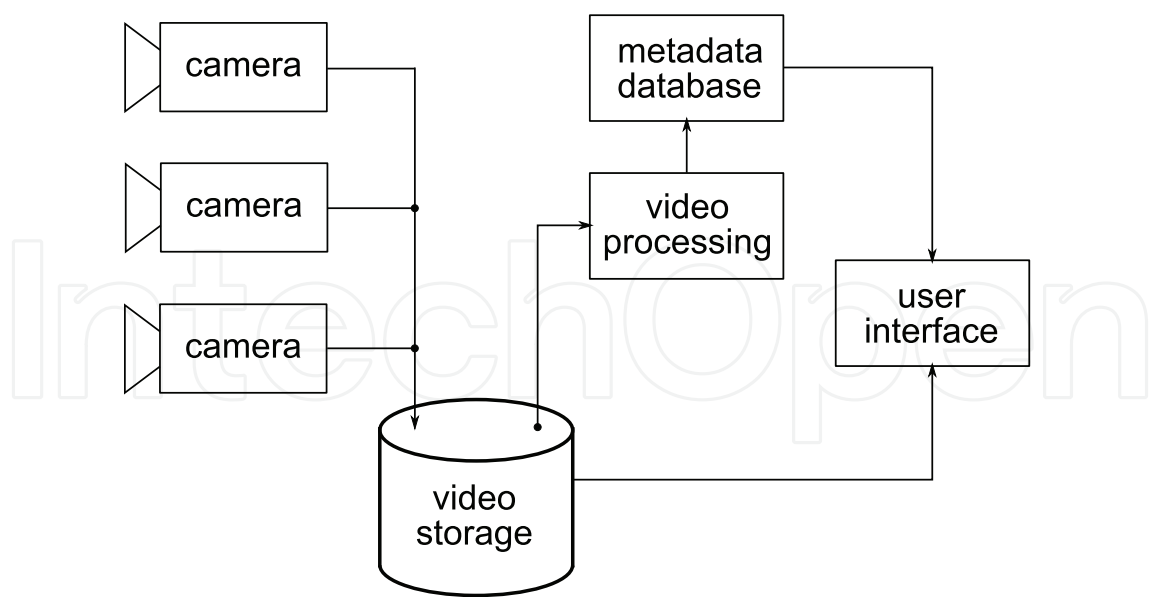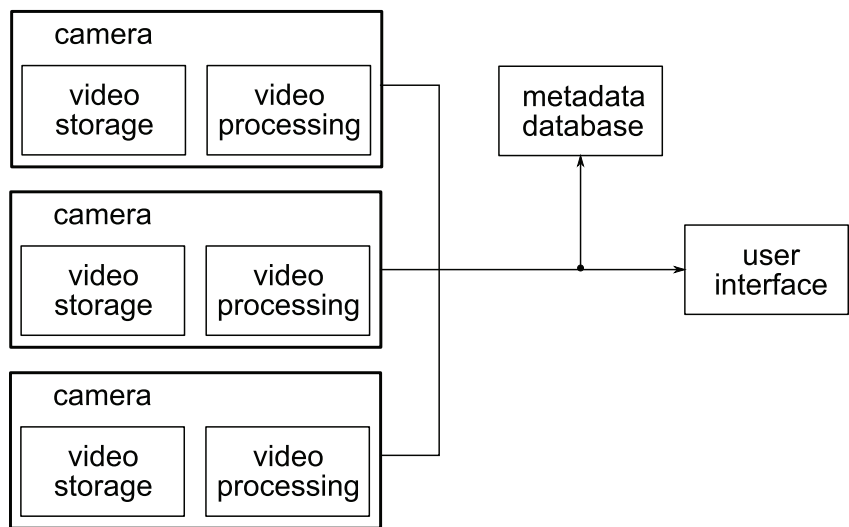
Fig. 15. Centralized architecture.

Fig. 16. Decentralized architecture with smart cameras.

A calibrated multicamera networks are useful to provide information on the distance between any object and camera. From this information 3D position of objects can be calculated. Also better models of objects (for example human) can be obtained. However, correct spatial information relies on the automatic matching between corresponding pixels in each image. This process is computationally expensive, for example pixels in low texture areas are very hard to match.

Imaging systems for assistive applications may not have the highest possible resolution. Often is more important to decrease bandwidth, also images with lower resolution are less difficult in terms of image processing. As an example we can mention system CARE (Safe private homes for elderly persons), wireless low-cost stereo vision system using cameras with resolution of 304×256 pixels, which is deployed to use in Germany and Finland Belbachir (2010).

11. Monitoring / video analysis

Monitoring and video analysis for assistive purposes generally means recognition of human activities and generation of action or alarm in case of unexpected behavior of the monitored person. The problem of human action recognition is quite complicated but with adequade choice of image processing methods is possible to find model of articulated non-rigid body. We aim to recognize five types of human daily activities: lying, sitting, standing, walking and other movements including transitions between sitting and standing or lying, and some leg movements when the human subject is sitting or lying – these movements are not assumed to be comparable to walking.

Generally it's possible to split problem of video analysis into following five points:

1. Extract some visual features from images or video sequence. Such features cover wide range of indicators from relatively simple (colour patterns, edges, histogram) to more complex (foreground/background estimation, segmentation, optical flow).

2. Tracking these features in video to obtain temporal sequence using Kalman filters or particle filters.

3. Classifying objects including variations within one class of objects (for example different human poses). For the object recognition can be used for example Haar classificators.

4. Recognition of spatio-temporal patterns like above mentioned lying, sitting, etc. For the tracking of human motion are used points or blobs based models – see Fig. 17.

5. Storing of obtained information as metadata linked to the original image or video data.

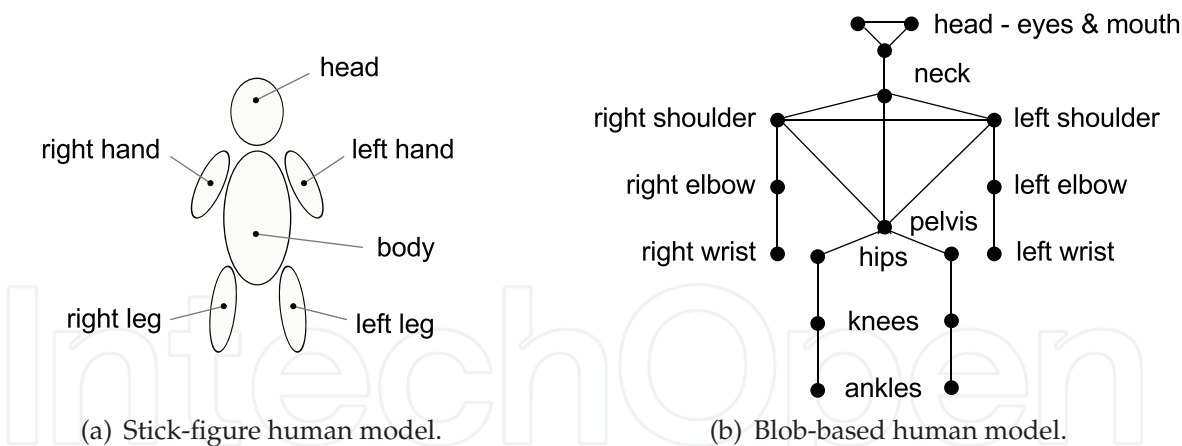(a) Stick-figure human model.   (b) Blob-based human model.

Fig. 17. Models of human body.

For the detection of various kind of objects could be successfully used Viola-Jones object detection framework Viola & Jones (2001), which is able to provide competitive object detection rates in real-time. It can be trained to detect a variety of object classes: human face, hand, upper body etc. During detection phase of the method a window of the target size is moved over the input image, and for each subsection of the image the Haar-like feature is calculated. (Simple Haar-like feature can be defined as the difference of the sum of pixels of areas inside the rectangle, which can be at any position and scale within the original image. Each feature type can indicate the existence or absence of certain characteristics in the image,

such as edges or changes in texture.) This difference is then compared to a learned threshold that separated non-object from objects. Beacuse such a Haar-like feature is only weak classifier, in the Viola Jones object detection framework the features are organized in so called *classifier cascade* to form a strong classifier.

For human activity or behavior recognition, most efforts have been concentrated on using state-space approaches. State space models have been widely used to predict, estimate, and detect signals over a large variety of applications. Most common used model is perhaps the Hidden Markov Model (HMM), which is a probabilistic technique for the study of discrete time series. Model defines each static posture as a state. These states are connected by certain probabilies. Any motion sequence as a composition of these static poses is considered a tour going through various states. Joint probabilies are computed through these tours, and the maximum value is selected as the criterion for classification of activities.

Another approach is to use the template matching technique to compare the feature extracted from the given image sequence to the pre-stored patterns during the recognition process. The advantage of using the template matching technique is its inexpensive computational cost; however, it is relatively sensitive to the variance of the movement.

The autonomous monitoring systems should be able to generate alarms. These system can be *rule based* (system is hard programmed) so that when condition or combination of conditions satisfies a certain rule, then apropriate alarm is triggered) or *learning based* – in this case system accumulates typically probability functions to represent what occurs frequently, when an outlier is detected an alarm can be triggered. More informations about automatic video processing can be found in Velastin (2009).

12. Safety & security

Fall is one of the most common causes of injury of seniors. Fractures healing due to falls is bad for old people and therefore is very important to prevent fals (for example by analysis of human balance) or generate an alert if fall happened. Fall-down incident normally occurs suddenly within approximate 0.45 to 0.85 seconds, see for example Chen (2010). When that happens, both the posture and shape of the victim change as shown in Fig. 18. It means that fall detection can be simply based on bounding box aspect ratio measurement.

For the identification of particular persons in the image may be successfully used face recognition based on the three main algorithms: PCA (Principal Component Analysis), LDA (Linear Discriminant Analysis), and EBMG (Elastic Bunch Graph Matching).

With PCA, probe image and training data of the same size must be normalized to line up the eyes and mouth of the subjects within the images. The PCA approach is then used to reduce the dimension of the data and decompose the face structure into orthogonal (uncorrelated) components known as eigenfaces. Each face image may be represented as a weighted sum of the eigenfaces; a probe image is compared against a training data by measuring the distance.

LDA is a statistical approach for classifying samples of unknown classes based on training samples with known classes. This technique aims to maximize between class (i.e. across users) variance and minimize within-class (i.e. within user) variance. When dealing with high dimensional face data, this technique faces the small sample size problem that arises

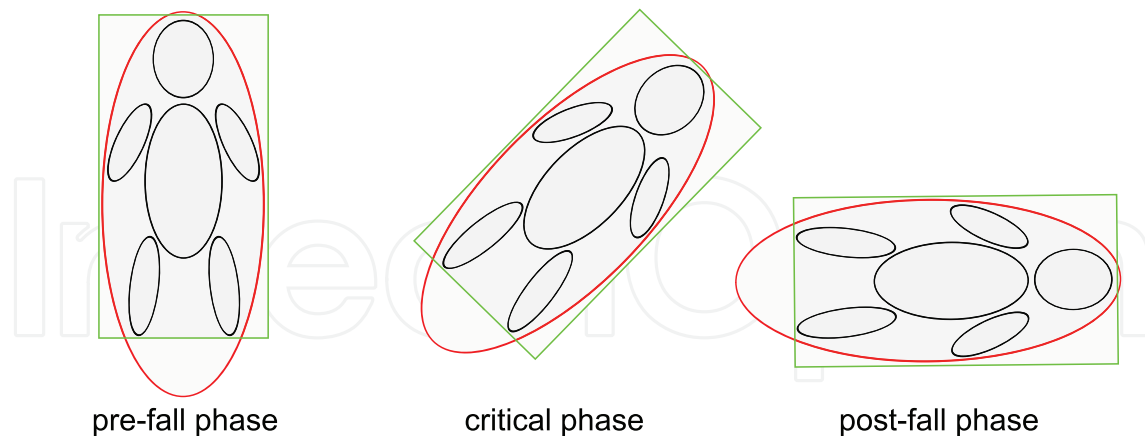pre-fall phase   critical phase   post-fall phase

Fig. 18. Phases of fall-down incident.

where there are small number of available training samples compared to the dimensionality of the sample space.

EMGM relies on the concept that real face images have many nonlinear characteristics that are not addressed by the linear analysis methods discussed earlier, such as variations in illumination, pose and expression. A Gabor wavelet transform creates a dynamic link architecture that projects the face onto an elastic grid. Then recognition is based on similarity of the Gabor filter response at circles around nodes of this grid. EMGM method is sometimes combined with PCA or LDA.

13. Combination with other sensors – smart home

The multimedia imaging systems and related technology, which we mentioned in the text, can be effectively utilized in smart homes. Smart home, for example see Fig. 19, is the place where are combined various kind of sensors in order to control people who live here. Outputs of those sensor are processed either automatically or generates alerts for supervising center.

The first of the additional sensors is RFID (Radio-frequency identification). The purpose of RFID is mainly for the fast identification of the persons in the database (i.e. know persons), for example for room access control or identification of the use of potentially hazardous things. Persons have a bracelet that includes a passive most probably unique RFID tag. The RFID readers are typically distributed at the doors of the rooms – when a person crosses a door, the RFID reader detects the bracelet and identifies the person who wears it. Besides to this identification functionality, each RFID reader controls the operation of the electronic lock of the door, allowing or not the access to the room depending on the identity of the person that is going to use it.

Other additional sensor, which works mainly locally as well as RFID, is PIR sensor. This sensor is useful particularly as a detector of presence – it is important in places where is not necessary to identify a specific person or in places where the presence of cameras might be annoying, such as a bathroom, where privacy of the person must be ensured.
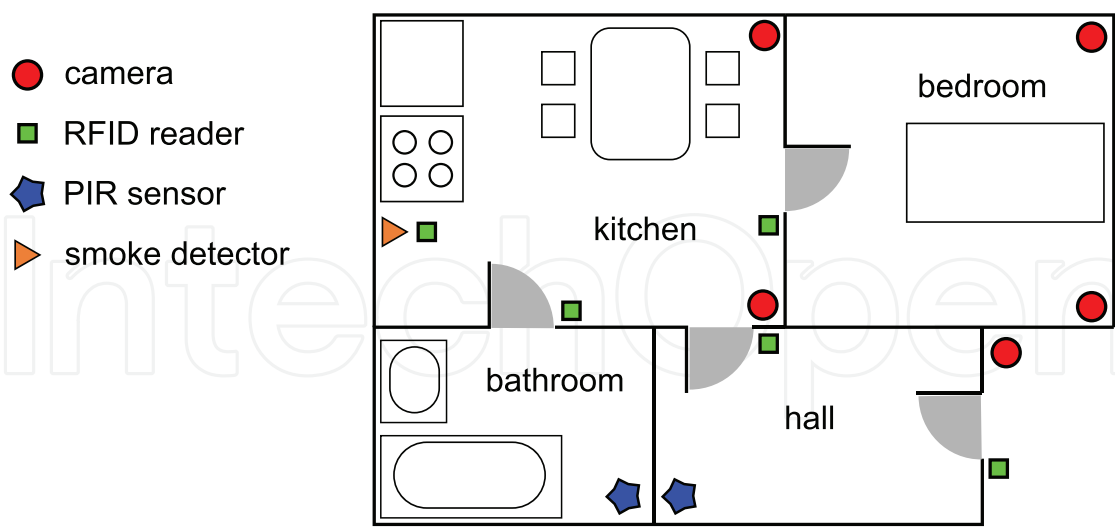
Fig. 19. Example of smart home.

Smoke detector is a typical representative of the group of detectors generating global alerts. Together with the gas detector are used in places where is risk of fire, gas leaks of even explosion.

The last part of smart home we would like to mention is inteligent bed. This bed integrate for example pressure sensors with ECG or other sensors. Simplest usage of pressure sensors in the inteligent bed is monitoring of daily activity of a person: bed-exit time atd. More advanced usage of pressure sensors is measurement of sit-stand movement – it is suitable method for early detection of decreasing mobility, which could help in the prevention of falls. In adition of measurement with pressure sensors is possible to use imaging system also – measured person perform assigned motion, during which is acquired model of human body and the positions of the key points of the skeleton are evaluated – see Fig. 20.

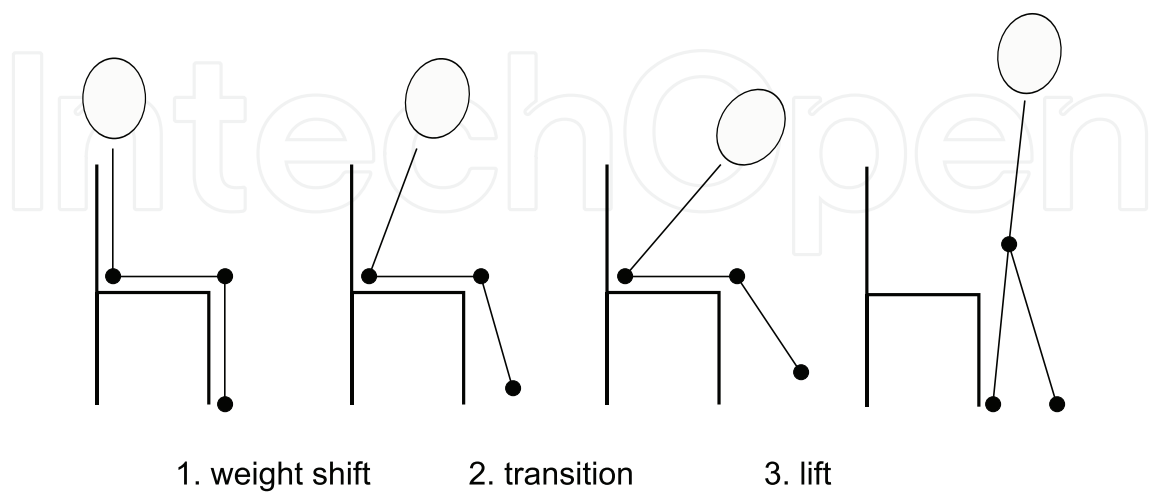1. weight shift   2. transition   3. lift

Fig. 20. Phases of sit-stant movement.

14. Conclusions

The Assistive Technology covers a broad range of different techniques and approaches streaming to the same aim – support and assistance for elderly and handicapped persons. It is a typical interdisciplinary field highly demanding not only towards designers and manufacturers but also towards users. In this chapter has been carried out general overwiev of applied multimedia imaging systems. The first part was especially a review of a wide range of available technology for image sensing, reproduction, transmission and compression. Knowledge of the properies of multimedia technology is necessary for the design of assistive application, which is devoted to second part. Processing of video or sequences of images is focused mainly to recognition of human activities, which is together with other sensors base of smart homes – facilities, which will play an increasingly important role in the assistance of our old or handicaped fellow citizens. We can also expect increase of usage of devices like smart phones etc. so data fusion with this kind of devices must be included in design od new assistive technology.

15. References

Lukac, R. (2009). *Single Sensor Imaging*, CRC Press, ISBN 978-1-4200-5452-1.

Fiete, R. D. (2010). *Modeling the Imaging Chain of Digital Cameras*, SPIE Press, ISBN 978-0-8194-8339-3.

Keelan, B. W. (2001). *Handbook of Image Quality*, Marcel Dekker, ISBN 0-8247-0770-2.

Gonzales, R. C. & Woods, R. E. (2008). *Digital Image Processing*, Pearson Prentice Hall, ISBN 978-0-13-168728-8.

Shi, Y. Q. & Sun, H. (2008). *Image and Video Compression for Multimedia Engineering*, CRC Press, ISBN 978-0-8493-7364-0.

Shi, Y. & Real, F. D. (2010). Smart Cameras: Fundamentals and Classification, *Smart Cameras*, Springer, ISBN 978-1-4419-0952-7

Aghajan, H. & Cavallaro, A. (2009). *Multi-Camera Networks, Principles and Applications*, Elsevier, ISBN 978-0-12-374633-7.

Senior, A. (2009). An Introduction to Automatic Video Surveillance, *Protecting Privacy in Video Surveillance*, ISBN 978-1-84882-301-3.

Belbachir, A. N & Lunden, T. & Hanák, P. & Markus, F. & B ottcher, M. & Mannersola, T. (2010). Biologically-inspired stereo vision for elderly safety at home, *Elektrotechnik und Informationstechnik*, Vol. 127, No. 7-8, ISSN 0932-383X.

Aggarwal, J. K. & Cai, Q. (1997). Human motion analysis: a review, *Nonrigid and Articulated Motion Workshop*, IEEE proceedings, pp. 90-102.

Velastin, S. A. (2009). CCTV Video Analytics: Recent Advances and Limitations, *Visual Infomatics: Bridging Research and Praice*, ISBN 978-3-642-05035-0.

Fan, Q. & Bobbitt, R. & Zhai, Y. & Yanagawa, A. & Pankanti, & S. Hampapur, A. (2009). Recognition of Repetitive Sequential Human Activity, *Computer Vision and Pattern Recognition*, ISBN 978-1-4244-3992-8

Villacorta, J. J. & Val, L. & Jimenez, I. & Izquierdo, A. (2010). Security System Technologies Applied to Ambient Assisted Living, *Knowledge Management, Information Systems, E-learning, and Sustainability Research*, pp. 389-394, ISBN: 978-3-642-16317-3.

Rougier, C. & St-Arnaud, A. & Rousseau, J. & Meunier, J. (2011). Video Surveillance for Fall Detection, *Video Surveillance*, Weiyao Lin (Ed.), ISBN 978-953-307-436-8.

Chen, Y. & Lin, Y. & Fang, W. (2010). A Video-based Human Fall Detection System for Smart Homes, *Journal of the Chinese Institute of Engineers*. Vol. 33, No. 5, pp. 681-690.

Scanaill, C. N. & Carew, S. & Barralon, P. & Noury, N. & Lyons, D. & Lyons, G. M. (2006). A Review of Approaches to Mobility Telemonitoring of the Elderly in Their Living Environment, *Annals of Biomedical Engineering*, Vol. 34, No. 4, pp. 547-563.

Arcelus, A. & Veledar, I. & Goubran, R. & Knoefel, F. & Sveistrup, H. & Bilodeau, M. (2011). Measurements of Sit-to-Stand Timing and Symmetry From Bed Pressure Sensors, *IEEE Transaction on Instrumentation and Measurement*, Vol. 60, No. 5, pp. 1732-1741.

Viola, P. & Jones, M. (2001). Robust Real-time Object Detection, *Second International Workshop on Statistical and Computational Theories of Vision – Modeling, Learning, Computing, and Sampling*.

Assistive Technologies

Edited by Dr. Fernando Auat Cheein

This book offers the reader new achievements within the Assistive Technology field made by worldwide experts, covering aspects such as assistive technology focused on teaching and education, mobility, communication and social interactivity, among others. Each chapter included in this book covers one particular aspect of Assistive Technology that invites the reader to know the recent advances made in order to bridge the gap in accessible technology for disabled or impaired individuals.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

INTECH
open science | open minds