National Institute of Justice (NIJ)
Sensor, Surveillance, and Biometric Technologies (SSBT)
Center of Excellence (CoE)

Camera Technology Survey of 3D Offerings and 2D Occlusions

December 3, 2012

Prepared by:

azimuth incorporated
1000 Technology Drive, Suite 3120
Fairmont, WV 26554
(304) 363-1162

NIJ SSBT CoE Operated By:

ManTech International Corporation

This project was supported by Award No. 2010-IJ-CX-K024, awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. The opinions, findings, and conclusions or recommendations expressed in this publication/program/exhibition are those of the author(s) and do not necessarily reflect those of the Department of Justice.

Contents

Executive Summary

Azimuth Incorporated, as a subcontractor of ManTech International Corporation, is supporting the National Institute of Justice (NIJ) Sensor, Surveillance, and Biometric Technologies (SSBT) Center of Excellence (CoE) program in the areas of test and evaluation of prototype biometric technologies and devices. This document describes the market survey and findings for 3D cameras and camcorders capable of face collections at 50 meters or greater, and an assessment of 2D camera background removal algorithms and techniques.

Capturing face images of uncooperative individuals at a distance of greater than 50 meters, in unconstrained environments is a very challenging task. Matching those collected faces in an automated fashion against poor legacy mug shot datasets is even more challenging. Given current capabilities of face matching algorithms, poorly collected probe images, and poor quality enrolled images, the resulting matching process performance is abysmal and worse in that the false matches consumes valuable resources to have a human to adjudicate false matches. The National Institute of Standards and Technology (NIST) has proven that face recognition can be very accurate with existing face matching algorithms, if the probe and dataset images are of excellent quality. Unfortunately, high resolution studio quality images can rarely, if ever, be collected covertly by surveillance cameras. Generally, at a distance unconstrained face image collections suffer due to insufficient resolution; insufficient/improper lighting, which causes hot spots or shadows on the face; complex image backgrounds, which may make face segmentation more difficult; collection camera angle or head rotation, which results in less than a full frontal face image collection; face obstructions caused by eyeglasses, hats, or hair; and atmospheric effects, which distort the image. A 3D camera or camcorder may mitigate or compensate face shadows, image background segmentation, and correcting the image to a full frontal face.

Commercially available 3D camera and camcorder products have made significant improvements, but may never be an optimal solution to long range face collections due to unique imaging system requirements for biometric collection and matching. Commercial products are produced for the mass market, price sensitive volume sales to individuals, where the constraints placed on face recognition systems are not designed into the imaging systems. Generally to meet face collection imaging system constraints, at 50 meters and greater, requires a high resolution, global shutter sensor that has very good visible frequency quantum efficiency, a minimum 300mm lens, and a large objective aperture to prevent image blur due to a slow shutter and/or subject movement. The large objective aperture gathers as much light as possible, and the lens system focuses the image on the sensor, so that the shutter can open and close again fast enough to prevent blur in the image collected while maintaining good exposure. To the CoE's knowledge, currently there exists no substitute for a large objective aperture in the imaging system, although there are research systems that take advantage of techniques such as synthetic and distributed aperture. Also, there are experimental lens systems that provide a large aperture through folding the lens optical path parallel to the sensor surface instead of perpendicular (normal) like traditional camera systems. An origami lens[(1)] would be much thinner or shorter than a traditional lens of similar focal length, which would make the camera system concealable. The Defense Advanced Research Projects Agency (DARPA) Military Imaging and Surveillance Technology – Long Range (MIST-LR)[(2)] research project could generate camera technologies useful to law enforcement surveillance, but would not likely result in a low cost commercial product. The survey of commercial 3D cameras and camcorders returned very few products capable of face collections at 50 meters or greater distance.

The 3D camcorders and cameras identified in this survey were not capable of meeting the law enforcement surveillance mission. The best 3D camcorder products this survey identified as potential face recognition image collection systems are the Sony DEV-3[(3)], JVC GS-TD1[(4)], and Panasonic HDC-10000[(5)]. There were no 3D cameras identified that were capable of collecting face images at 50 meters or greater that provided the resolution and focal length required to provide for 90 pixels between pupil centers. Neither the JVC nor the Panasonic are ideal solutions for collecting face images at a distance. The Sony DEV-3 participated in a Spring 2012 face collection effort at West Virginia University (WVU) as part of the evaluation of the NIJ SSBT CoE SVI binocular. The current hardware version of the SVI binocular does not have an auto focus capability nor does it provide image stabilization. The Sony binocular hardware has auto focus, image stabilization, and an ability to collect 3D video and was used to compare image quality and match performance to the SVI binocular. As an example, the Sony DEV-3 collected a 3D video recording of a person at 50 meters and the face image was extracted. The extracted face image inter-pupillary pixel count was 22, which is approximately one fourth the recommended 90 pixels. Another data point to consider is at the same distance using a Canon 5D Mark II camera body (21 megapixel sensor) with an 800mm fixed focal length lens, the inter-pupillary pixel count was 159. The primary concern is that the cameras will not provide enough pixels center of pupil to center of pupil to allow automated face recognition of images collected. The specifications provided by the manufactures for each of the 3 identified 3D face collection systems do not provide enough technical data to calculate the number of inter-pupillary pixels they would provide at 50 meters. Also, due to the suspected poor performance (due to low sensor pixels counts and short focal lengths) of collecting 3D face images for automated matching, it is not recommended that these systems are tested for this application. Commercial 3D cameras and camcorders are not currently capable of filling the needs of law enforcement for long distance (50 meters and greater) face collections and automated matching. In addition to 3D camcorders and cameras this survey also reviewed products and research associated with face segmentation.

The most promising research for segmentation is being conducted at Stanford University in the area of computational photography. A specific example is a paper written by Vabhav Vaish titled, "Synthetic Aperture Focusing using a Shear-Warp Factorization of the Viewing Transform." A web link to a video demonstrating the capability is provided in Appendix A[(8)]. This research provides possible face segmentation by background and foreground removal. In essence, it offers the capability to provide individual face segmentation from the image, background and foreground, so long as there are no other faces on the same focal plane. Another interesting facet of the research is the ability to rotate the focal plane in any orientation within the capture volume rather than just normal to the sensor plane. In addition to providing face segmentation, it can also provide an ability to see through objects that partially occlude the face of interest. These capabilities have interesting applications to capturing surveillance images for face segmentation and automated matching. The current research implementation is a multi-camera system which is computationally intensive. This research may produce commercial products in the future as did the light field research at Stanford leading to the Lytro camera.

Scope

To help frame the CoE's evaluation of Long-Range 3D Facial Recognition technology, there is a need to conduct market surveys and research activities into Commercial-Off-the-Shelf (COTS) 3D (i.e., stereoscopic) cameras and video recording devices. Advances in facial recognition and background occlusion technologies have continued to move forward, reportedly making particular advances in the last 18 to 24 months. The objective of this task will be to perform a market survey of 3D recording devices (both still capture and video) detailing the available devices and vendors, and to investigate the capabilities of 2D background removal technologies. These objectives will provide a fresh view of the current state of available technologies which perform similar functions as the Stereo-Vision Imaging Incorporated (SVI) prototype, as well as provide feedback to R&D efforts at SVI.

Background

Biometrics is an automated method of identifying a person based on physiological or behavioral characteristics. Biometric technologies are becoming the foundation of many law enforcement mission sets requiring identity management. This research is focused on identifying image capture systems (cameras and camcorders) and image post-processing necessary to remove or segment an individual's face from the background in the image. This captured and segmented face image can then be provided to feature extraction and matching algorithms. The following paragraphs provide background on image capture and post capture processing to segment the face from the image background information.

At the heart of any face matching capability is the quality of the face images that are captured, both the enrolled image (gallery) and live captured image (probe). Controlled image capture environments with cooperative subjects generally result in images where high quality features can be extracted, greatly improving matching algorithm performance. The NIST challenge events for face matching algorithms is where this notional result is well documented. In the "real" world, the collection environment is not controlled and the enrollee may not cooperate, or even be aware, during the collection process. Some of the real world collection problems encountered during long range face collections are extreme lighting conditions and angles that produce shadows on the face (artificial indoor and daylight outdoor), subject motion, pose angle of the face, facial expressions, occlusions (glasses, hats, scarves, and facial hair), outdoor environmental constraints such as precipitation (rain, fog, and snow), dust, and air temperature changes between the camera and collection subject which result in image distortions. All of these collection conditions individually will reduce the quality of the images collected and therefore the matching performance of face biometric matching algorithms. Collectively, these real world constraints can cause face matching to provide no benefit at all in identifying individuals and in the worst case, waste law enforcement resources investigating false matches and conversely allow true matches to remain unknown. This research of surveillance system technologies and computational image processing algorithms identified COTS products that can mitigate some of these real world collection constraints.

This investigation and research included a survey of current COTS products in image capture and processing; research efforts in industry, academics, and Government laboratories; and fundamental enabling technologies in the areas of sensors, optics, and image processing. The primary purpose of the imaging system is to acquire images where high quality features can be extracted for matching. The imaging system is comprised of the optics, image sensor, and in most cases pre-processing of the image which occurs internal to the camera or camcorder.

Many of the current biometric systems attempt to mitigate poor imaging system performance by providing a quality algorithm that measures the quality of the collected sample. If the sample is of poor quality the system will acquire another sample, many times without intervention from the operator. This process continues until the system acquires an image of good quality or the system "times out" and returns a Failure To Acquire (FTA). The face modality does not have a universally accepted quality algorithm such as the NIST Fingerprint Image Quality (NFIQ) for the fingerprint modality. In comparison, each face recognition vendor provides their own quality algorithm which causes face matching efforts to be less effective in large scale National collection and matching systems. This is particularly true of legacy mug shot datasets where image quality can vary from very poor to excellent. Commercial face matching algorithms have different and particular sensitivities to pose angle, expressions, and shadows. A commercial matching and quality algorithm paired together can provide better match performance than one paired with a generic quality algorithm. The assumption is that vendors knowing the sensitivities of their matching algorithms can better tune the associated quality algorithm. Collecting a high quality face image is a function of collector training, imaging system, lighting and background control, and post image capture enhancement software.

Tradeoffs that affect optimization of the face image collection system are optics and sensor characteristics; this tradeoff is most often made based on cost considerations. As an example, the lens system may consist of cheap plastic elements rather than glass. In addition to advances in sensors and lenses, software processing of images has made major improvements to the performance of image capture systems. The ability of each technology to separately improve image capture quality has matured and stagnated. Recent research and product advances in image acquisition seem to be a convergence of sensor, lens, and software engineering. The most striking example of this convergence is in the field of computational or light field photography. Optimizing the overall performance of a face biometric identification system also includes removal of the image background from the individual face of interest and providing the segmented face image to the feature extraction and matching algorithms. The first step is image analysis to determine if it contains a face and then segmentation of the face from the background is required. Face segmentation from the background of the image is a critical step in achieving an overall good match performance for the system (details are provided in the "COTS Background Removal Algorithms" section).

After acquiring a face image, background segmentation is an essential step in automated face matching systems. Providing face images to the feature extraction and matching algorithms requires that a face image is found in the collected image and that the face is segmented from the background of the image. Current segmentation algorithms are based in one of two popular methods, discontinuity and similarity. Discontinuity is the abrupt changes in gray level, such as edges in an image. Similarity is an approach to segmentation that partitions an image into regions that are similar according to a set of predefined criteria. In addition to face segmentation the following paragraphs provides information into other areas of research being conducted in imaging systems that could have applications in 50 meter and greater distance face collection and recognition systems.

The Federal Government, industry, and academics have invested much effort and resources in novel optical systems that change classical design paradigms. The Defense Advanced Research Projects Agency (DARPA) Optical Non-Redundant Aperture Generalized Sensors (MONTAGE) under the Microsystems Technology Office (MTO), MDO, Compressive and Optical MONTAGE Photography (COMP-I)[6], and many more programs advanced new techniques that would benefit a face recognition system such as folded optics, adaptive optics, digital super resolution, computational photography, and compressive imaging. Face recognition systems rely on image sensors and optics to acquire a face image and also image processing algorithms to provide a useful face image to the feature extraction and matching algorithms. These enabling technologies of sensors, optics, and image processing are always improving and therefore could increase the performance of face recognition systems. A limited review of enabling imaging technology research is described in the paragraphs that follow.

Sensors products from Kodak, DALSA Semiconductor, Fairchild Imaging, Rockwell, and others have application to face recognition. Sensors are generally divided into major classes, CCDs and CMOS, both with advantages and disadvantages. In general the sensor characteristics, cost not considered, most desired for a facial recognition system are: High quantum efficiencies (sensitivity or low light gathering) - A CCD has better Qe – This characteristic manifest itself in the ability to collect images in low light and could reduce blur due to a shorter integration time on the sensor. Global shutter is better than rolling because of image distortion caused by rolling shutter collections. Resolution and pixel size are generally a tradeoff. The more resolution and larger the pixel is generally creates better images. Optical sensors are always providing a constant performance improvement in the following areas:

- Charge Coupled Device (CCD) or Complementary Metal Oxide Semiconductor (CMOS)
- Global or Rolling Shutter
- Color Filter Type (RGB or YMC)/Spectral Response
- Resolution
- Image Size & Aspect Ratio/Format
- Pixel Pitch/Size
- Photo Sensitivity/Quantum Efficiency – within the spectrum of interest
- Dynamic Range/Bit Depth
- Read-Out Noise/Amplifier Noise
- Dark Current Noise
- Maximum Pixel Charge / Pixel Well Depth
- Anti-blooming Control Technique
- Maximum Read Time/Data Rate

Optical systems continue to improve due to computer modeling but a slower advance in materials this area has stagnated. Improved performance is still possible due to new coatings which reduce internal lens reflections. Other more novel research is a folded optics project funded by DARPA that would provide a smaller and lighter weight lens for each sensor.

Image processing algorithms, and the hardware where they execute, are improving and providing capabilities that were not possible only a few years ago. Not only are the capabilities improving but the cost to implement these capabilities is also decreasing. Image processing algorithms are very diverse in functionality and are a topic unto itself. For example, algorithms exist to improve dynamic range of an image, autofocus, image stabilization, and low light imaging. Research identified lens and imaging processing solutions from many vendors to include those from Intel, DxO Labs, CDM Optics, Distant Focus Corporation, Plainsight Systems, Digital Optics, and others.

Survey Task Market Survey of COTS 3D Cameras and Camcorders

This survey attempted to identify potential commercially available candidate devices that are capable of performing 3D capture at a distance up to 50 meters. The survey results specify the vendors offering the devices and investigate the technical specifications of the devices, output formats, and feasibility of integration of the technology into existing facial recognition systems. The specific survey findings are provided in the paragraphs that follow.

COTS 3D Cameras & Camcorders

The requirement to collect face images at 50 meters places resolution constraints on the imaging system. For mobile biometric collections, a face resolution of 90 pixels between centers of pupils is recommended by NIST best practices. This resolution recommendation is intended for collections at about 1 - 2 meters. However to ensure matching accuracy, this recommendation is used for camcorder and camera analysis in this report for collections at 50 meters.

The ability of each face collection product to provide 90 pixels between pupils would require an analysis of each product. The variables to analyze for each system would be the horizontal resolution of the sensor, the horizontal Field of View (FOV), and the distance from the camera to the subject face being collected. The accepted mean for distance between the centers of pupils for an average adult is 63 mm, 5th percentile 55mm and 95th percentile 70mm. The field of view would change based on the focal length of the lens. The principal difficulties in obtaining high quality images at ranges of 50 meters or greater for this application are 1) the focal length required of the optical system, 2) the ability of the lens system to collect enough light to quickly capture an image, and 3) the ability of the imaging sensor to provide an inter-pupil resolution of 90 pixels or greater at distances of at least 50 meters.

Long focal length optics is required to provide the face resolution required for biometric identification. In general a focal length of about 600 mm is needed for facial recognition at a 50 meter range. The practical implication is that existing optical systems with the required focal length are incompatible with the small, lightweight, easily transportable systems generally needed by Law Enforcement. The long focal length lens makes it difficult to quickly find the individual of interest in a small field of view and hold the face in the collection frame during collection. In addition, during collections a long focal length lens is subject to image blur due to movement and vibration of the camera and/or subject. Addressing the second issue of sufficient light, there does not appear to be a substitute for a large aperture on the objective of the optic system. The longer the lens must stay open to integrate an image on the sensor focal plane, the more susceptible the system is to blur in the face image collected.

The 3D camera and camcorders identified in this survey were not capable of meeting the law enforcement surveillance mission application at a minimum of 50meters. The sensors were rolling shutter rather than global and the lenses focal lengths were not long enough to collect the necessary inter pupil number of pixels. Also, the primary aperture for the lenses was small on the camcorders and cameras causing longer exposure times which results in blur and smearing for less than ideal lighting conditions.

The 3D Camera and Camcorder products that most closely met the long range face surveillance requirements were: the Sony DEV-3, JVC GS-TD1, and Panasonic HDC-10000. Although these were the most suitable to collecting face images at 50 meters or greater none are appropriate for the task. As an example, the Sony DEV-3 collected a 3D video recording of a person at 50 meters and the face image was extracted from the video. The extracted face image inter-pupillary pixel count was 22 which is one-fourth the NIST recommended 90 pixels. Another data point to consider is at the same distance using a Canon 5D Mark II camera body (21 megapixel sensor) with an 800mm fixed focal length lens the inter-pupillary pixel count was 159, not quite twice the recommended number of pixels. The primary concern is that the cameras will not provide enough pixels center of pupil to center of pupil to allow reliable automated face recognition of images collected. The specifications provided by the manufactures for each of the 3 identified 3D face collection systems do not provide enough technical data to calculate the number of inter-pupillary pixels they would provide at 50 meters. Also, due to the suspected poor performance (due to lower sensor pixel counts and short focal length lens) of collecting 3D face images for automated matching, it is not recommended that the systems are tested for this application.

The underlying survey assumption is that a 3D collected face image will provide better match performance in real world conditions than a 2D face image that is collected under similar conditions. In addition, nearly all criminal legacy face datasets are 2D, which would require that the 3D images be converted to 2D prior to feature extraction. Considering current product 3D collection sensor and optics capability in the visible spectrum, other technologies such as LiDAR may prove superior at creating 3D face images. Imaging systems in the InfraRed (IR) spectrum suffer from many of the same limitations of face collection systems in the visible spectrum, primarily resolution and contrast. Also consider that advances in near and medium wave infrared active illumination systems can make night and poor weather face collections more reliable. Infrared sensors are generally more expensive, lower resolution, and require more power for active cooling. Active illumination, particularly at long range, requires a lot of power and precludes covert collections. Research is being conducted on multi spectrum fusion systems and multi-spectral sensors, for example SRI International Sarnoff[7], but these are not COTS products designed for large scale production and sale to the general public.

There are products beginning to emerge from an area of photography research called computational photography. The Government has funded an extensive research portfolio for computational photography with researchers at government laboratories, academics, and industry. The first product sold to the general public is a “light field” camera manufactured by Lytro[9]; Raytrix[10] also sells industrial computational cameras. The Lytro camera is the result of Ph.D. work in light field camera research conducted by Ren Ng at Stanford University. The Lytro camera captures not only the quantity of photons, as a traditional camera does, but also a direction from which it came. In general, traditional cameras record a scalar quantity and the Lytro camera records a vector quantity. The Lytro product only scratches the surface of the possibilities computational photography offers. In addition to being able to refocus an image during post image processing, other possible capabilities include total image focus (from closest to most distant objects in an image), changing camera and scene perspective post image capture, and 3D stereo images collected with a single lens and sensor. Lytro has a planned firmware update that will allow the camera to collect 3D images. In effect computational photography will allow personalized images to be created thru post processing after initial capture. These technologies promise vastly improved camera surveillance systems and improved face recognition performance in real world environments.

Survey Task Assessment of 2D Camera Background Removal Techniques

This survey performed an assessment of techniques used to isolate faces from the background of captured images. The assessment examined the current capabilities of algorithms and techniques used to perform face-extraction from a 2D image by removing or ignoring non-face portions of the image. While the primary focus of this assessment was on commercially available solutions, an effort was made to include promising technologies still in Research and Development. The depth and breadth of the assessment was primarily an internet search and dependent on the information availability from technical papers and vendors.

COTS Background Removal Algorithms

Image face segmentation is an essential first step in most automated face matching systems. Providing face images to the feature extraction and matching algorithms requires that a face image is found in the collected image and that the face is segmented and extracted from the image. Major commercial camera manufacturers now integrate face finding algorithms, and even blink detection, into their products. These manufactures algorithms are proprietary and trade secrets but discussed here are the general approaches to face finding and background segmentation.

Two major techniques are required to face segmentation where there is a controlled or an uncontrolled background. Controlled static backgrounds, where only faces are expected to appear, are the simplest to segment and extract. Each live collected image is subtracted from a stored static background image of the same scene using a Haar Cascade Face Detector (Viola-Jones algorithm). This technique can be used in fixed surveillance locations such as airports or access control. In the case of uncontrolled backgrounds segmentation algorithms can be based in one of two methods, discontinuity and similarity. Discontinuity is the abrupt changes in gray level, such as edges in an image. Typically, discontinuity utilizes some form of an edge detector to identify rapid changes in brightness or contrast. The canny edge detector algorithm is an example of edge finding that is widely used in image processing. An example of similarity is an approach to segmentation that partitions an image into regions that are similar according to a set of predefined criteria. These two techniques partition the image into smaller areas using a sliding block that are sequenced to ensure that the entire image is examined. Each block is evaluated and determined to either contain or not contain a face. Principal Component Analysis (PCA) can then be used to determine if faces are present in each of the image blocks. PCA method uses a faces dataset represented by a set of eigenvectors and a threshold value. The probe image is projected into a low dimensional feature space that is spanned by the eigenvectors of a set of test faces. The resulting principal components are then compared. The threshold is used to determine if the image contains a face. An example of an uncontrolled background might be a dash camera in a police car.

Other methods of face finding and segmentation rely on face color or motion detected between video frames. Finding faces using color information is unreliable due to changing lighting conditions and the variation in human skin color. The motion between frames method assumes that motion from frame to frame of a video is a face. Depending on the application this can be a very flawed method of face finding. Although poor at finding faces it does have application to face tracking. These individual techniques can be combined to provide more accurate face finding and segmentation. A current research segmentation technique that utilizes computational photography is reviewed in the following section.

Algorithm R&D Survey

A promising research project for segmentation is being conducted at Stanford University in the area of computational photography. A specific example is a paper written by Vabhav Vaish titled, "Synthetic Aperture Focusing using a Shear-Warp Factorization of the Viewing Transform." A web link to a video demonstrating the capability is provided in Appendix A[(8)]. This research provides the future possibility of background and foreground removal. In essence it provides the capability to provide individual face segmentation from the image, background and foreground, so long as there are no other faces on the same focal plane. Another interesting facet of the research is the ability to rotate the focal plane in any orientation within the capture volume. The image plane of current commercial cameras is always parallel to the sensor plane. In addition to providing face segmentation, it can also provide an ability to see through objects that partially occlude the face of interest. These capabilities have interesting applications to capturing surveillance images for face segmentation and automated matching. The project hardware implementation is a multi-camera system which is computationally intensive. This research may produce commercial products in the future, as did the light field research leading to the Lytro camera.

Conclusion

The 3D camcorder products this survey reviewed as potential face recognition image collection systems are the Sony DEV-3, JVC GS-TD1, and Panasonic HDC-10000 (specifications provided in Appendix B). There were no 3D cameras or camcorders identified that were capable of collecting face images at 50 meters or greater that provided the resolution and focal length required to provide 90 pixel between pupil centers. Neither the JVC nor the Panasonic are ideal solutions for collecting face images at a distance. The Sony DEV-3 is currently collecting face image data at WVU as part of the evaluation of the NIJ SSBT CoE SVI binocular. The Sony binocular hardware has auto focus, image stabilization, and an ability to collect 3D video and will be used to compare image quality and match performance to the SVI binocular. Commercial 3D cameras and camcorders are not currently capable of filling the needs of law enforcement for long distance, 50 meters and greater, face collections and automated matching.

The most promising research area associated with image face segmentation is in computational photography. This is a fundamental shift in the way in which cameras process and record light. Current generation cameras capture and store the number of photons only, whereas next generation cameras will also know where they came from within the capture volume of the scene (photons as scalar quantities verses vectors). In essence, instead of a human viewable image on a camera sensor plane the image is created inside the processor. This camera technology shift will allow face segmentation from images in ways which are currently impossible using today's cameras. These future technologies have interesting applications to capturing surveillance images for face segmentation and automated matching. This research will provide new camera products and capabilities but as always first generation systems will likely be less capable and expensive.

APPENDIX A - Product and Research Web links

[1] 'Origami lens' slims high resolution cameras, University of California - San Diego

http://www.jacobsschool.ucsd.edu/news/news_releases/release.sfe?id=617

Origami Lens Slims High Resolution Cameras, http://www.jacobsschool.ucsd.edu

[2] Defense Advanced Research Projects Agency (DARPA) Military Imaging and Surveillance Technology – Long Range (MIST-lR)

(3), (4), & (5) are product specifications and found in Appendix B

[6] Multiple Optical Non-re-dundant Aperture Generalized Sensors (MONTAGE) Program - MDO Primary Participants: DARPA, University of Arizona (Lead organization), University of California at San Diego, Massachusetts Institute of Technology, CDM Optics, Inc., Distant Focus Corporation, and CDM Optics, Inc.

[7] SRI International Sarnoff - http://www.sarnoff.com/markets/defense-security

[8]Synthetic Aperture Focusing using a Shear-Warp Factorization of the Viewing Transform, Vabhav Vaish

Video demonstrating the capability is at: http://graphics.stanford.edu/papers/shear-warp/a3diss.avi

[9] Lytro - http://www.lytro.com/#

[10] Raytrix - http://www.raytrix.de/

Appendix B - 3D Camcorder Specifications

[3] SONY DEV-3 Specifications

Sony DEV – 3 Basic Specifications

Weights and Measurements

- Dimensions (Approx.) : Approx. 10.63inch x 2 5/ 8inch x 3.47inch (219mm x 155mm x88mm)
- Weight (Approx.) : Approx. 2.65lbs (body only)

Exposure System

- Minimum Illumination : Standard: 11 lux (1/60 Shutter Speed:60i, 1/50 Shutter Speed:50i)

Optics/Lens

- Aperture : F1.8-3.4
- Focal Length (35mm equivalent) : Movie Mode: 57.9-660mm(2D)/34.4-344mm(3D), Photo Mode: 29.8-660mm(16:9),27.4-606mm(4:3)
- Lens Stabilization : Optical SteadyShot™ Image stabilization w/ Active mode (wide (3D)/ Wide to Tele (2D))
- Lens Type : G Lens™
- Magnification : PHOTO:x0.5-10x(2D:DigitalZoomOFF)/x0.3-x7.5(2D:DigitalZoomON); MOVIE:x0.9-x10(16:9)/x0.5-x5.4(4:3)
- Minimum Focus Distance : 3D: Approx. 80cm (Wide), Approx. 7.5m (Tele); 2D: Approx. 1cm (Wide), Approx. 80cm (Tele), Approx. 25cm (Tele Macro)

Advanced Features

- Image Stabilization : Optical SteadyShot™ Image stabilization w/ Active mode (wide (3D) / Wide to Tele (2D))

Imaging Sensor

- Imaging Sensor : 1/4" (4.5mm) Back-illuminated "Exmor R" CMOS Sensor
- Pixel Gross : 2x Approx. 4200K pixels (3D)/ Approx. 4200K pixels (2D)
- Processor : BIONZ™ Image processor

Recording

- Audio Format : Dolby® Digital 2ch Stereo

- Media Type : Internal Flash Memory; Memory Stick PRO Duo ™ (Mark 2); Memory Stick PRO-HG Duo ™; SD/ SDHC/ SDXC Memory Card (Class 4 or Higher)
- Microphone/Speaker : Built-in Stereo Microphone Monoral Speaker
- Still Image Max Effective Resolution : 7.1 megapixels
- Still Image Mode : 3D Still Image: Not available / 2D Still Image: JPEG
- Still Image Size(Photo Mode) : 7.1 megapixels 4:3 (3072x2304) 5.3 megapixels 16:9 (3072x1728) 1.9 megapixels 4:3 (1600x1200) 307K pixels 4:3 (640x480)
- Video Format : 3D HD: MPEG4-MVC/ H.264 AVCHD 2.0 format compatible (1080/60i); 2D HD:MPEG4-AVC/H.264 AVCHD™ format ver.2.0 compatible (1080/60p/60i original format)
- Video Mode : 3D HD: Approx. 28Mbps; 2D HD PS: Approx. 28Mbps, FX: Approx. 24Mbps, FH: Approx. 17Mbps, HQ: Approx. 9Mbps, LP: Approx. 5Mbps
- Video Resolution : 3D HD: 2x 1920x1080/60i; 2D HD: 1920x1080/60p (PS), 1920x1080/60i (FX,FH), 1440x1080/60i (HQ,LP)
- Video Signal : NTSC

Convenience Features

- Multiple Language Display : English; Simplified English; Japanese, Traditional Chinese, Simplified Chinese, French, Italian, German, Spanish, Dutch, Russian, Romanian, Protugese, Canadian French, Latin American Spanish, Korean, Turkish, Polish, Czech, Hungarian, Indonesian, Maly

Interface

- A/V Remote Terminal : Video / S Video / Audio / Component Out / Remote
- BRAVIA® Sync™ : Yes
- DC IN : Yes
- HD Output : HDMI Out (mini); Component Video Out (Component A/V cable) (sold separately))
- HDMI Terminal : Yes (mini)
- Headphone Jack : Stereo Minijack
- Memory Card Slot : Memory Stick PRO Duo™ and SD/ SDHC/ SDXC media compatible
- Microphone Input : Stereo Minijack
- SD Output : Composite Video Out (A/V connection cable (supplied)); Component Video Out (Component A/V cable (sold separately)); S Video Out (A/V connection cable with S Video (sold separately))
- USB Port(s) : TypeA, mini-AB/USB2.0 Hi-speed (mass-storage/MTP)

Viewfinder

- Diopter Adjustment : Yes (-3.5 - +3.5)
- Eye relief : 14.3mm
- Field of View : 35.6degrees (16:9)
- Interpupilary distance : 55-75mm
- Type : 2x 0.45-inch (1.1cm) Wide (16:9); 1227k dots (852x3(RGB)x480)

(4) JVC GS-TD1

JVC GS-TD1 Basic Specifications

Optics

Sensor - 2x 3.32 Megapixel 1/4.1" CMOS Sensor

Zoom - Optical: 10x & Digital: 200x [1]

Recording

System – NTSC

Recording Media - Internal Flash Memory 64GB SD/SDHC/SDXC

Recording Time – Internal Flash Memory

3D Mode (AVCHD): TXP: 16 hour, 30 minutes; TSP: 23 hours, 30 minutes
3D Mode (MP4 MVC): THR: 8 hours, 10 minutes; TSR: 12 hours, 30 minutes
2D Mode (AVCHD): UXP: 11 hours, 40 minutes; EP [3]: 59 hours

Video Format - **High Definition MPEG AVC/H.264** 1920 x 1080

Display

Type	-	LCD
Size	-	3.5"
Touch Screen	-	Yes

Features

Image Stabilization	Optical
Built in Mic	Yes
Built in Speaker	Yes
Tripod Mount	1/4"

Input/Output Connectors		
	Inputs	1x 1/8" (3.5mm) Stereo Mini
	Outputs	1x USB 2.0 1x HDMI C (Mini) 1x 1/8" (3.5mm) Stereo Mini 1x A/V
General		
System Requirements		Windows XP, Vista, 7
Battery		Rechargeable Lithium-Ion Battery Pack
Dimensions (WxHxD)		4.02 x 2.52 x 7.32" / 102 x 64 x 186 mm
Weight		1.49 lb / 675 g

(5) Panasonic HDC-10000

Panasonic HDC-10000 Basic Specifications

Image Device 2x 3MOS (3x CMOS each) 1/4.1" sensors Total Pixels 9.15 megapixels (3.05 megapixels x 3) Effective Pixels **Motion Image**
2D: 6.57 megapixels (2.19 megapixels x 3) [16:9]
3D: 6.21 megapixels (2.07 megapixels x 3) [16:9]

Still Image
7.08 megapixels (2.36 megapixels x 3) [3:2]
6.57 megapixels (2.19 megapixels x 3) [16:9]
6.87 megapixels (2.29 megapixels x 3) [4:3] Lens 2x f/1.5 lenses with 12x zoom in 2D; 10x zoom in 3D Focal Length
2D: 2.84-34.1 mm
3D: 2.84-28.4 mm 35mm Equivalent Focal Length **Motion Image**
2D: 29.8-368.8 mm [16:9]
3D: 32.0-320.0 mm [16:9]

Still Image
2D: 29.8-369.0 mm [3:2], 29.8-368.8 mm [16:9], 31.0-372.0 mm [4:3]
3D: 32.0-320.0 mm [16:9] Aperture 2D: f/1.5 (wide) to f/2.8 (tele)
3D: f/1.5 (wide) to f/2.7 (tele) Zoom **Optical**
2D: 12x
3D: 10x

Intelligent
2D: 23x

Digital
2D: 30x / 120x Horizontal Resolution *Not specified by manufacturer* Sensitivity *Not specified by manufacturer* Standard Subject Illumination 1400 lux Minimum Illumination 5 lux (iA mode ON: 1/30 second shutter) Minimum Object Distance (M.O.D.) 2D: Approx. 47" (1.2 m) full zoom range
Approx. 1.4" (3.5 cm) wide range macro

3D: Approx. 2.5 m (all zoom range)
Approx. 18" (45 cm) wide range macro Focus Auto / Manual White Balance Auto / White Set / 3200K / 5600K / Ach / Bch Iris Auto / Manual Vertical Smear *Not specified by manufacturer* Signal System 3D: 1080i60, 1080p30, 1080p24

2D: 1080p60, 1080i60, 1080p30, 1080p24 [1] Built-in Filters No LCD Monitor 3.48" (8.83 cm) 16:9 3D (glasses-free) ILCD monitor (1,152,000 dots) Viewfinder 0.45" (11.5 mm) 16:9 EVF (approx. 1,226,880 dots) Shutter Speed **Auto**

Slow Shutter On
60p, 60i, 30p mode: 1/30 to 1/8000 second
24p mode: 1/24 to 1/8000 second

Auto Slow Shutter Off
60p, 60i, 30p mode: 1/60 to 1/8000 second
24p mode: 1/48 to 1/8000 second Gain Selection *Not specified by manufacturer* Signal-to-Noise Ratio *Not specified by manufacturer* Image Stabilization 2D: HYBRID O.I.S. +
3D: POWER O.I.S. Manual Controls Rings: Zoom, focus, iris
Dial: Convergence Memory Card Slot 2x SD card (SD up to 2 GB, SDHC up to 32 GB, and SDXC up to 64 GB)
Recording Format AVCHD 2.0 standard (AVCHD 3D / AVCHD Progressive) compliant Compression Technology 2D: MPEG-4 AVC / H.264
3D: MPEG-4 MVC / H.264 Bit Rates MVC (3D) & 1080p60 (2D): 28 Mb/s
PH: 24 Mb/s
HA: 17 Mb/s
HE: 5 Mb/s Maximum Recording Time *With 64 GB storage*
MVC (3D) & 1080p60 (2D): 5 hours 15 minutes
PH: 6 hours
HA: 8 hours 30 minutes
HE: 27 hours 30 minutes Still Capture Recording Format: JPEG (DCF / Exif2.2), MPO (CIPA-007-2009)
Aspect Ratio: 16:9
Image Size: 3 Mp [JPEG], (2304 x 1296), 2.1 Mp [JPEG], (1920 x 1080), 2.1 Mp [MPO + JPEG], (1920 x 1080)
Thumbnail Display 20 thumbnails/page, 9 thumbnails/page, 1 thumbnail/page Audio Dynamic Range *Not specified by manufacturer* Audio Signal Format Dolby Digital 5.1-channel / 2-channel, Linear PCM 2-channel
48 kHz / 16-bit Audio Bitrates Dolby Digital: 384 kb/s (5.1-channel), 256 kb/s (2-channel)
Linear PCM : 1.5 Mb/s (2-channel) Built-in Speaker Dynamic type (round) Input and Output Connectors **XLR Inputs**
2x 3-pin (INPUT 1, INPUT 2)
LINE: 0 dBu, MIC: -50 dBu / -60 dBu (selectable in menu)

HDMI Output
1x Type A
Outputs 1080p, 1080i, 480p & 576p video and Dolby Digital & Linear PCM audio

Headphone Output
1x stereo mini jack (3.5mm diameter)

USB 2.0
1x USB Mini

AV Multi Output
Proprietary connector to component/composite/stereo audio (6x) RCAs Power Requirements 7.2V (battery) / 12V (AC adaptor) Power Consumption Maximum 16.7W (recording) Operating Temperature 32 to 104°F (0 to 40°C) Operating Humidity 10-80% (no condensation) Dimensions (WxHxD) Approx. 5.7 x 7.7 x 13.8" (14.5 x 19.5 x 35.0 cm) Weight Approx. 3.5 lb (1.6 kg) System Requirements For HD Writer XE 1.0 software (Windows-only):

OS
Windows 7 (32-bit/64-bit): Starter/Home Basic/Home Premium/Professional/Ultimate
Windows Vista (32-bit): Home Basic/Home Premium/Business/Ultimate SP1/SP2
Windows XP (32-bit): SP3

CPU
Intel Pentium 4 2.8 GHz or higher
When using playback function and MPEG-2 conversion function, Intel Core 2 Duo 2.16 GHz or higher

Operation on an upgraded OS is not guaranteed. When 2 or more USB devices are connected to a PC, or when devices are connected though USB hubs or by using extension cables, proper operation is not guaranteed. Operation on an OS other than the one pre-installed is not guaranteed. Not compatible with multi-boot environments or multi-monitor environments. Not guaranteed on Windows XP Media Center Edition, Tablet PC Edition, Windows Vista Enterprise, or Windows 7 Enterprise. Available with administrator account in Windows XP; available with administrator account or standard account in Windows Vista/Windows 7 (installation and uninstallation require an administrator account).

APPENDIX C - Acronyms and Abbreviations

Acronyms and Abbreviations

Acronym	Definition
AFIS	Automated Fingerprint Identification System
AIS	Automated Information Systems
ANSI	American National Standards Institute
ARGUS-IS	Autonomous Real-time Ground Ubiquitous Surveillance-Imaging System
CCD	Charge Coupled Device
CMOS	Complementary Metal Oxide Semiconductor
CoE	Center of Excellence
COMP-I	Compressive and Optical MONTAGE Photography
COTS	Commercial Off-The-Shelf
CPU	Central Processing Unit
DARPA	Defense Advanced Research Projects Agency
DOC	Department of Commerce
DOJ	Department of Justice
DSP	Digital Signal Processor
EBTS	Electronic Biometric Transmission Specification
EFTS	Electronic Fingerprint Transmission Specification
FAR	False Acceptance Rate
FBI	Federal Bureau of Investigation
FOV	Field of View
FPGA	Field Programmable Gate Array
FRR	False Rejection Rate
FTA	Failure to Acquire
FTE	Failure to Enroll
GIMP	GNU Image Manipulation Program
GOTS	Government Off-The-Shelf
IAFIS	Integrated Automated Fingerprint Identification System
JPEG	Joint Photographic Experts Group
MONTAGE	Multiple Optical Non-redundant Aperture Generalized Sensors
NIJ	National Institute of Justice
NIST	National Institute of Standards and Technology
NISTIR	NIST Interagency Reports
R&D	Research and Development
RISC	Reduced Instruction Set Computer

S&T	Science and Technology
SSBT	Sensor, Surveillance, and Biometric Technologies
T&E	Test and Evaluation
XCBF	XML Common Biometric Format