

Augmented Reality and Remote Sensing: Using Multi-Spectrum to Exhibit Our Physical Environment

Lingli ZHU, Juha SUOMALAINEN, Eero SALMINEN, Juha HYYPPÄ, Harri KAARTINEN, Arttu JULIN, and Hannu HYYPPÄ, Finland

Key words: Remote Sensing, Augmented Reality, Multispectral camera, NDVI, Object Detection, PCA, Image Matching, Similarity Measurement Android Mobile Application Platform

ABSTRACT

In this paper, we propose an AR technology in which our surroundings can be visually enhanced by superposing common video with information from multispectral bands such as red and infrared channels. The camera used in this study is a NIR camera with a spectral range of 600nm-975nm, which is commonly used for vegetation study. We utilize images acquired from the red and infrared channels to calculate the information about e.g. the vegetation NDVI and a customized object class and put them into a library. An Android smartphone based video camera is used to record the real environment, in the predominant scene, augmented information from the library can be retrieved and superposed on it.

Augmented Reality and Remote Sensing: Using Multi-Spectrum to Exhibit Our Physical Environment

Lingli ZHU, Juha SUOMALAINEN, Eero SALMINEN, Juha HYYPPÄ, Harri KAARTINEN, Arttu JULIN, and Hannu HYYPPÄ, Finland

INTRODUCTION

Augmented Reality (AR) is a technology to combine or mix the view of the real environment with additional, virtual content that is presented through computer graphics (Augmented Reality, 2017). It has been used in education, entertainment, GIS, media arts, psychology, robotics, surgery and urban planning. With the help of advanced AR technology, the information about the surrounding real world becomes interactive and digitally manipulable. It can bring out the components of the digital world into a person's perceived real world. However, the visible spectrum is limited to the visible light, approximately in the wavelength from 390nm to 700nm. Human eyes and brain can distinguish the pure colors and a mix of multiple wavelengths in such spectral range (Human Eye, 2017). However, in our physical world, the electromagnetic spectrum (EMS) covers a wide range, typically from Gamma rays, x-rays, ultraviolet, visible light, infrared light, microwaves and Radio waves, between 1pm to 100Mm (Electromagnetic spectrum, 2017). We can utilize remote sensing sensors to explore information which is out of human vision.

Remote sensing sensors measure electromagnetic radiation and interact with the atmosphere and the objects to acquire information including not only the distance between the sensor and the object, but also the direction, intensity, wavelength content, and polarization of electromagnetic radiation. In different EMS, objects exhibit different characteristics. For example, remote sensor working in green channel is typically applied for water body measurement due to less absorption by water body in that channel, while exploring the vegetation, red and near infrared lights are applied: to calculate the Normalized Difference Vegetation Index (NDVI), to classify different objects, to find the health status of the vegetation and tree species. Such information can't be directly acquired by human vision.

In previous study, a synergy to visualize physical and virtual city environments by integrating AR with photogrammetry has been addressed by Portales et al. (2010). The authors introduced a low-cost outdoor mobile AR application to integrate buildings of different urban spaces. High-accuracy 3D photo-models derived from close-range photogrammetry are integrated in real (physical) urban worlds. The augmented environment that is presented herein requires for visualization a see-through video head mounted display (HMD), whereas user's movement navigation is achieved in the real world with the help of an inertial navigation sensor. AR technology applied for Virtual training of parts assembly, which has been introduced by Horejsi (2015). It uses a conventional web camera to shoot a referential workplace with a worker. The proposed software solution processes the web camera image data and adds virtual 3D model instructions to the real image. The final image is presented on a monitor placed in front of the worker. Mura et al. (2015) integrated AR with sensing device for manual assembly workstations to reduce human errors in performing the operations. An

Augmented Reality and Remote Sensing: Using Multi-Spectrum to Exhibit Our Physical Environment (9029)
Lingli Zhu, Juha Suomalainen, Eero Salminen, Juha Hyypä, Harri Kaartinen, Arttu Julin and Hannu Hyypä (Finland)

FIG Working Week 2017

Surveying the world of tomorrow - From digitalisation to augmented reality
Helsinki, Finland, May 29–June 2, 2017

augmented reality (AR) equipment used to give to the worker the necessary information about the correct assembly sequence and to alert him in case of errors. The force sensor is placed under the workbench and it is used to monitor the assembly process by collecting force and torque data with respect to an XYZ reference system. Two AR devices have been tested in this application: a video-mixing spatial display and an optical see-through apparatus. Recently, AR has been widely used in education field. Foster (2016) developed an AR system for visually impaired students for the lectures. Weng et al. (2016) presents an AR system for Biology science education. During the lessons, dedicated stereoscopic and photo-realistic views were presented and also assert something such as a specific attribute, a rule or a function, applies to all the objects of the same kind, thus facilitating students in noticing, memorizing and understanding Biology concepts.

This paper is aiming for using the multi-spectral sensors to enhance our physical environment in which acquired information is beyond the range of our vision. We utilize a multispectral camera to obtain images from red and near-infrared bands. Information about vegetation is investigated, for instance, the NDVI values, detecting known and unknown objects from a scene and so on. Such information is stored in a library file. When a video camera (e.g. a smartphone) records the scene, a searching is conducted from the library file. When finding the matched object, the image from library overlaps on the video image to show up more information about the current object. In the near future, when multispectral camera becomes available in our smart phone, synchronized information from video camera and multispectral camera will make our AR technology easy to be implemented.

MATERIALS

The materials used in this paper include a Near-Infrared (NIR) hyperspectral camera, a smart phone with video camera and development platform, Matlab software. The model of the camera is Ximea ‘MQ022HG-IM-SM5X5-NIR’. The specifications of the camera are as follows:

Resolution: 409x217 pixels
Spectral range: 600-875nm
Sensor active area: 25 Bands
Dimensions (without lens) WxHxD: 26x26x31 mm
Frame rates: up to 170 cubes/s
Weight: 32 g

A smartphone ‘Samsung Galax Note Edge’ with Android operating system was used for development platform. The device is powered by a 2.7GHz quad-core Qualcomm Snadragon 805 processor, 3GB of RAM and has 64GB of internal storage. It features a 16P camera, heart rate monitor and 3000mAh battery.

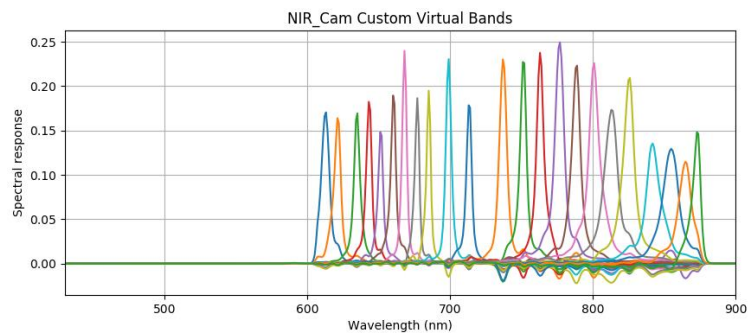


Figure 1. NIR camera; spectrum bands of the NIR camera; a smart phone.

METHODS

In this study, we concentrate on utilizing the advantages from multispectral data, especially in infrared bands, which are beyond human vision. It can help people understand the physical environment better. The multispectral camera works in the wavelength of 600nm-875nm, in which vegetation is highly sensitive. Some indicators/ parameters such as their healthy status or tree species can be obtained from these channels. Three cases are studied in this paper. They are i) to obtain the NDVI value of vegetation. NDVI images can allow people to see additional information on health status of leaves and vegetation; ii) to highlight and detect known objects. For example, it can help people to spot lost objects or mushrooms when those show up in a scene; iii) to highlight and alarm about object not fitting spectrally in a scenery. For example, this can be used to spot manmade objects in natural environment.

Image registration technology is conducted in our application. The method of structure similarity is highlighted. This method contains three primary measurements: luminance, contrast and structure. The development was implemented in two platforms. One is Matlab, another is Android smart phone developer platform. The details are as follows:

1. NDVI value acquisition

NDVI is a robust index which can be used to evaluate the presence and vigor of vegetation. Live green plants absorb efficiently solar radiation in the photosynthetically active radiation (PAR) spectral region which coarsely overlaps with the spectral range of human vision (NDVI, 2017).

Augmented Reality and Remote Sensing: Using Multi-Spectrum to Exhibit Our Physical Environment (9029)
Lingli Zhu, Juha Suomalainen, Eero Salminen, Juha Hyypä, Harri Kaartinen, Arttu Julin and Hannu Hyypä (Finland)

FIG Working Week 2017

Surveying the world of tomorrow - From digitalisation to augmented reality

Helsinki, Finland, May 29–June 2, 2017

Outside of this region and especially in the near-infrared (NIR) the vegetation cells have evolved to reflect and re-emit solar radiation in the near-infrared spectral region. A strong absorption at these wavelengths would only result in overheating the plant and possibly damaging the tissues. Hence, most organic objects appear relatively bright (>0.50 reflectance) in the NIR. Furthermore, the photosynthesis mechanisms in healthy green vegetation causes such objects to have appear very dark especially in red spectral region (<0.05 reflectance).

The NDVI exploits this difference between red and NIR response. With NDVI, the observed intensities in red and NIR bands are converted it to a single NDVI value describing health and organic nature of the target by using the following equation:

$$NDVI = \frac{NIR - RED}{NIR + RED}$$

Where NIR and RED are the observed intensities in the NIR and red regions of the spectrum ($\sim 800\text{nm}$ and $\sim 650\text{nm}$). The NDVI is a robust self-normalizing index and for coarse visualization of the NDVI the radiometric calibration of the sensor is not absolutely necessary but it is possible to achieve reasonable results with any digital numbers produced by the camera. For more advanced analysis calibration and conversion of observations to radiance or reflectance factor units is recommended.

2. Object detection-finding known objects

Normal human vision observes targets using three colors: red, green, and blue. This set of 3 spectral bands allows people to identify objects by their colors. The human sight is limited to these exact colors and number of bands due to evolutionary reasons, but when using camera technology there is no reason stick with this spectral range or such a low number of spectral bands. With our hyperspectral camera we take simultaneously measurement on 25 spectral bands in the spectral range from 600nm to 875nm. The human brain interprets “colors of objects“ using the natural RGB observation and uses those to detect and identify objects in view. Using hyperspectral cameras we can very similarly form the more advanced 25-bands spectrum of an object to describe its “color” in very high detail. If we compare this observed spectrum, with the spectrum of a to-be-searched object, we can scan the image for objects with that spectrum.

A simple way to evaluate similarity of the two spectra (x and y in equation below) is by using algorithm called Spectral Angle Mapper (SAM):

$$SAM = \arccos \left(\frac{\sum_{i=1}^{N_{bands}} x_i y_i}{\sqrt{\sum_{i=1}^{N_{bands}} x_i^2} \sqrt{\sum_{i=1}^{N_{bands}} y_i^2}} \right) \quad [1]$$

If the spectrum of the -to-be-found target is known, the hyperspectral image can be scanned for pixels of such spectrum and if similarity better than threshold is detected the user can be alerted of the observation.

3. Object detection-finding unknown objects

In the case above, the hyperspectral information was used to detect specific known objects. However this methodology is not possible if the spectrum of the to-be-found object is not known. In such case it is still possible to use the hyperspectral information to highlight potentially interesting objects in the scene. Especially in the natural environment, the spectrum of majority of objects in the scene can usually be described for most part using just a few spectral component shapes. Meanwhile anomalies, such as man-made objects lost in the nature, often have radically different spectral shape. This difference in spectral composition can be used to spot and highlight anomalous objects even if their spectrum is unknown.

4. PCA component analysis

A common method to determine the spectral component shapes dominating in the imagery is Principal Component Analysis (PCA). In PCA workflow, the spectra of the pixels are analyzed to determine the dominant spectral shapes presenting in imagery. The presence of these spectral component shapes is measured in each spectrum. Using these data, it is possible to form a mask highlighting all anomalous objects in the scene by setting a threshold on how many dominant spectral shape components are ignored.

5. Image registration

Image registration is necessary when measurements were taken in different time, different viewpoints or in different coordinate systems. Image registration makes different measurements comparable. Common image registration methods include point-based matching (e.g. intensity), feature-based matching (e.g. line, edge, and area), or similarity measurement. It can be interactive operation or automated operation. In our task, we conducted image matching using the automatic method for similarity measurement. Two datasets were utilized: one is a library file, which consists of the predefined images; another is the video images taken by the users. An image similarity measure quantifies the degree of similarity between intensity patterns in two images. Common image similarity measures include cross-correlation, mutual information, sum of squared intensity differences, and ratio image uniformity. At present, in the field of video image registration, the typical method is to measure structural similarity (SSIM). Structural information is the idea that the pixels have strong inter-dependencies. These dependencies carry important information about the structure of the objects in the visual scene. The SSIM formula is based on three comparison measurements between the samples: luminance, contrast and structure. This method improved the traditional methods such as peak signal-to-noise ratio (PSNR) and mean squared error (MSE).

The SSIM index is calculated on various windows of an image. The measure between two windows is (Wang et al. (2003)):

$$SSIM(x, y) = [I_M(x, y)]^{\alpha M} \cdot \prod_{j=1}^M [c_j(x, y)]^{\beta j} [s_j(x, y)]^{\gamma j} \quad [2]$$

Where $x = \{x_i | i=1, 2, \dots, N\}$ and $y = \{y_i | i=1, 2, \dots, N\}$ are two discrete non-negative signals that have been aligned with each other; $I_M(x, y)$ is the luminance comparison, computed only at scale M ; $c_j(x, y)$ is the contrast comparison at the j -th scale; $s_j(x, y)$ is the structure comparison calculated at the j -th scale; α, β, γ are parameters to define the relative importance of the three components.

When an object defined in the library appears in a scene, the similarity is measured. A threshold is set to determine whether they are the same objects.

6. Added information overlapping on video based on Android Smartphone Platform

A Samsung smart phone is based on the Android operation system. Android gives a powerful platform for creating apps. Android also gives you tools for creating apps that take advantage of the hardware capabilities available on each device. It gives the developers as much control as he wants over his UI on different device types.

The Android framework includes support for various cameras and camera features available (Class: *android.hardware.camera2* or *Camera*) on devices, allowing users to capture pictures and videos in their applications. A series of declarations of request permission is needed to use the devices, features and actions from the camera. To capture or stream images from a camera device, the application must first create a camera capture session with a set of output Surfaces for use with the camera device, with *createCaptureSession*. Each Surface has to be pre-configured with an appropriate size and format (if applicable) to match the sizes and formats available from the camera device. A target Surface can be obtained from a variety of classes, including *SurfaceView*, *SurfaceTexture* via *Surface(SurfaceTexture)*, *MediaCodec*, *MediaRecorder*, *Allocation*, and *ImageReader*. When recording a video, retrieval is processed to check whether the target surface is matching the image from library. When the matching condition is met, the library image shows up on the phone.

DEMONSTRATION OF THE RESULTS

After the implementation of our developed AR system, we are able to overlap multi-spectral information from a library on a video image. The aim is to demonstrate the operability of our method. The test data was a home plant which has many wide leaves in a good condition. The multi-spectral camera recorded imagery in red and near infrared bands. The plant was recorded in different forms: the plant, and the plant with green pins. We have achieved the following results i) obtained its NDVI image; ii) recognized known spectral object in the scene; iii) found abnormal objects in a scene by the PCA of the image; iv) acceptable image matching results; v) overlapped images shown on the smart phone.

Figure 2. A photo and a NDVI visualization of an indoor plant. In this visualization we can see how: (1.) the NDVI effectively masks the plant from the background, (2.) greatly reduces the effect of shadows in the plant structure (3.) highlights a dead leaf in the middle of the plant, and (4.) reveals the veins in the leaves on high detail.

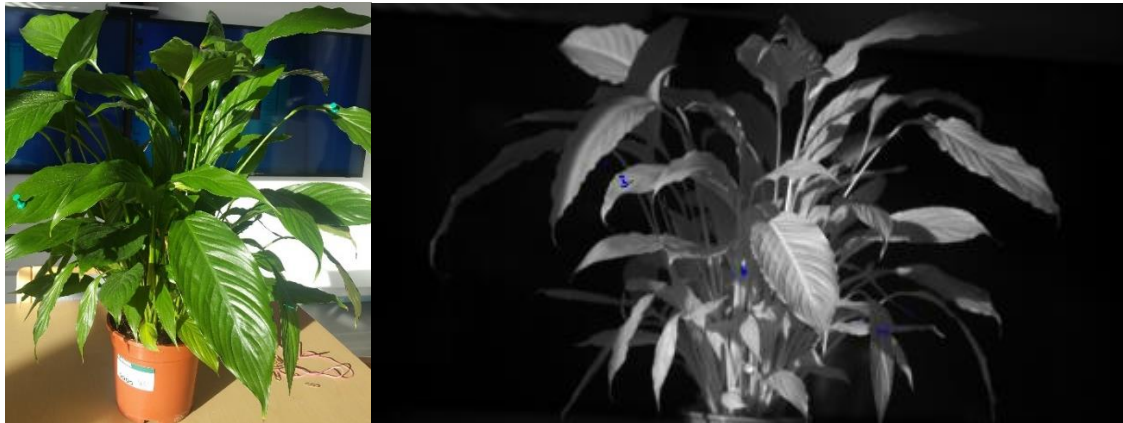


Figure 3. A photo and highlighted NIR image of the SAM detection experiment. Some green plastic pins were stuck to the leaves of the plant. For human eye, detecting all pins would have been a challenging task. From hyperspectral image, the spectrum of one pin was picked. This reference spectrum was compared to all pixels in the image using SAM algorithm. By setting threshold to $SAM < 0.10$ (radians) we were able to highlight the pins in the scene.



Figure 4. A photo and image of anomalous objects produced with unsupervised PCA filtering. To bring out the anomalous objects, a PCA transformation was done to the image, four most dominant spectral components were removed, and the PCA image was reverse-transformed back to normal hyperspectral spectral data cube. In the resulting data cube, the manmade anomalous objects (plastic pins and rubbers) are clearly showing up as they stand out spectrally from the majority of the scene. a). Plant with green pins; b) Plant with plastic buckets; c) PCA without any component removal; d) Investigate 12 top most dominant component contributions. The fifth component clearly shows the green pins; e) The seventh component shows the rubbers.



Augmented Reality and Remote Sensing: Using Multi-Spectrum to Exhibit Our Physical Environment (9029)
Lingli Zhu, Juha Suomalainen, Eero Salminen, Juha Hyypä, Harri Kaartinen, Arttu Julin and Hannu Hyypä (Finland)

FIG Working Week 2017

Surveying the world of tomorrow - From digitalisation to augmented reality

Helsinki, Finland, May 29–June 2, 2017

(e)

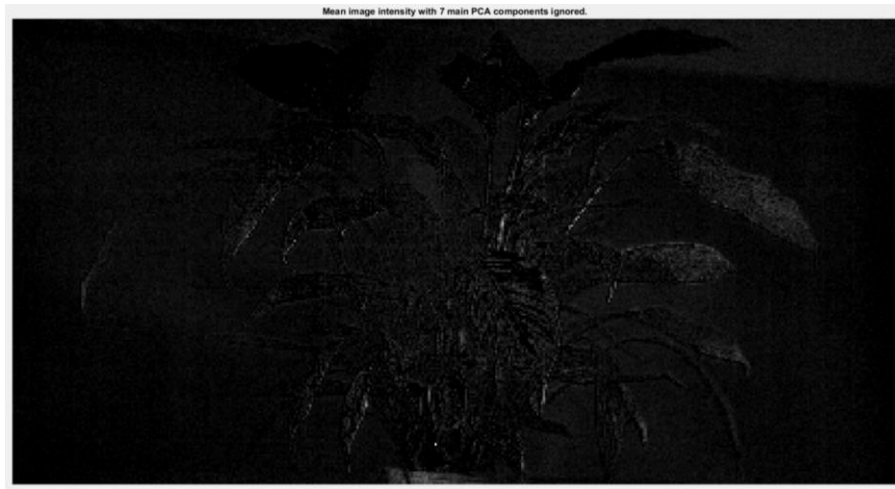


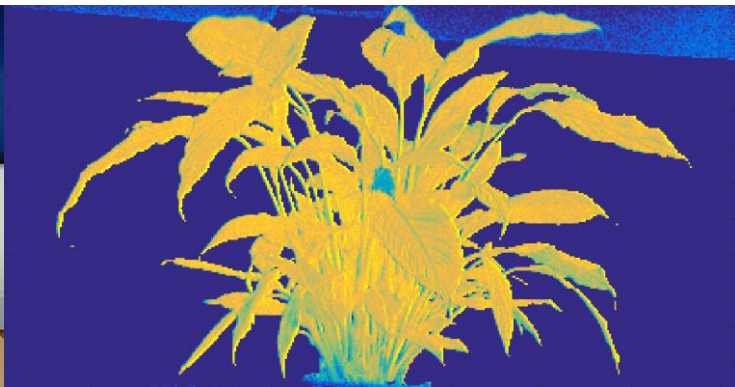
Figure 5. Image Registration and overlapped with the video image.

Plant's NDVI image was stored in a library. The library file was retrieved and the NDVI image was matched to the video image which was taken by a smart phone. The similarity measurement between two images was over 60%. Due to a bit different views from two images, the matching was not perfect, but they have showed a relatively high similarity. Finally the NDVI image was showed up on the video camera. (a) A plant image; (b) NDVI image of the plant; (c) NDVI image mask; (d) A video image; (e) NDVI image matching to video image; (f) Overlapped images when matching condition meets.

(a)



(b)



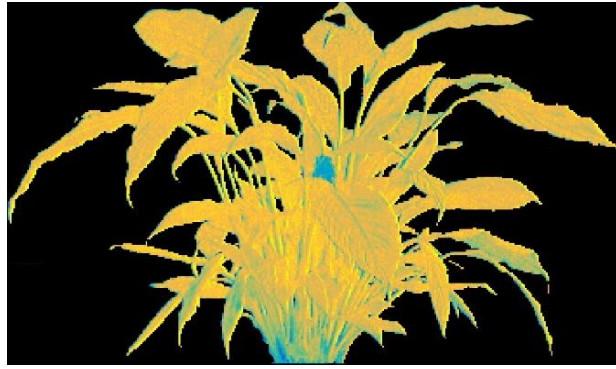
(c)

Augmented Reality and Remote Sensing: Using Multi-Spectrum to Exhibit Our Physical Environment (9029)
Lingli Zhu, Juha Suomalainen, Eero Salminen, Juha Hyypä, Harri Kaartinen, Arttu Julin and Hannu Hyypä (Finland)

FIG Working Week 2017

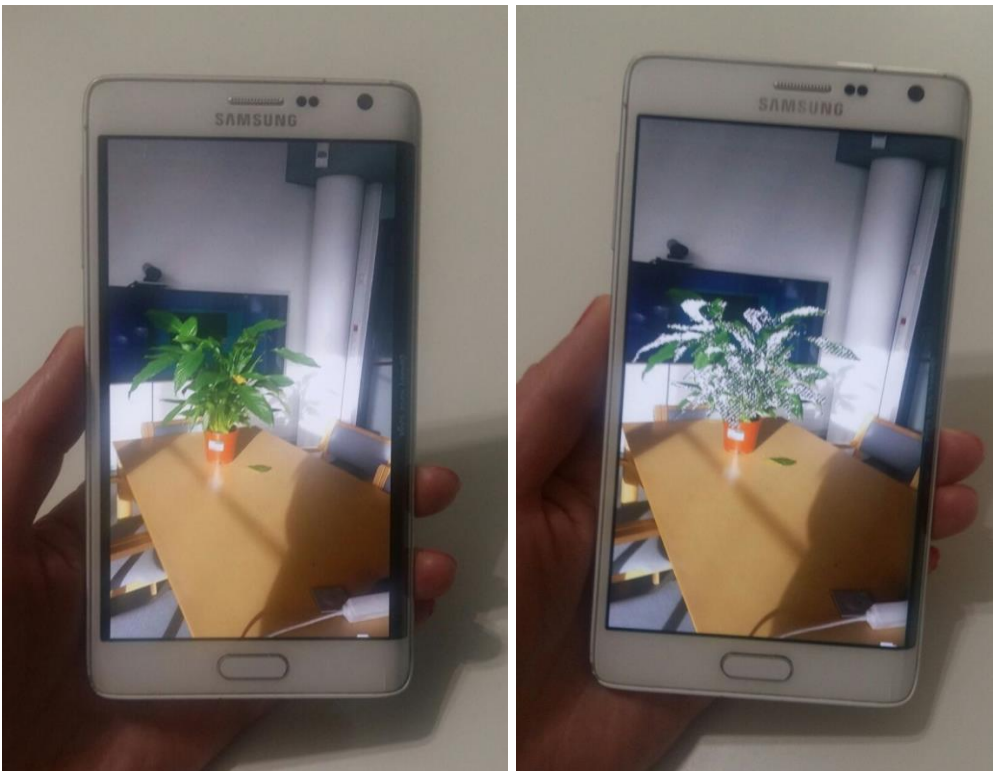
Surveying the world of tomorrow - From digitalisation to augmented reality

Helsinki, Finland, May 29–June 2, 2017



(d)

(e)



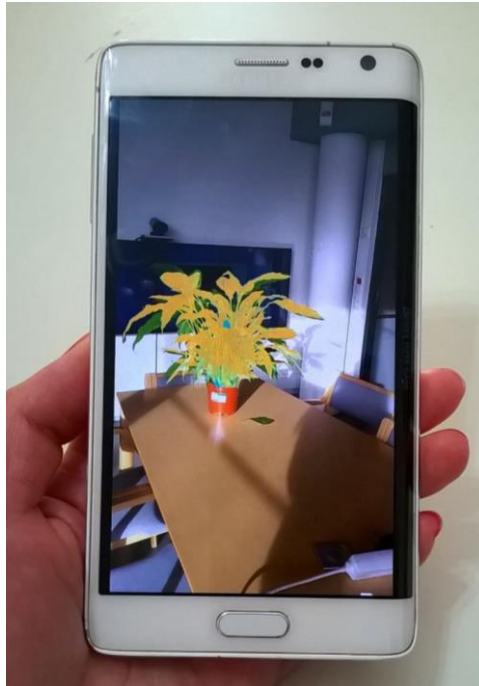
(f)

Augmented Reality and Remote Sensing: Using Multi-Spectrum to Exhibit Our Physical Environment (9029)
Lingli Zhu, Juha Suomalainen, Eero Salminen, Juha Hyypä, Harri Kaartinen, Arttu Julin and Hannu Hyypä (Finland)

FIG Working Week 2017

Surveying the world of tomorrow - From digitalisation to augmented reality

Helsinki, Finland, May 29–June 2, 2017



We have demonstrated the procedure of the experiment. We have utilized multi-spectral images to obtain the NDVI image of a plant. We have shown the possibilities to detect known spectral objects from a scene and also it showed that unknown objects can be revealed by PCA component analysis. These two applications can be extended to other similar applications. For example, when we know the spectrum of a mushroom, we set it as known object in a scene. Then when mushrooms emerged in a scene, an indicator can show the locations of the mushrooms. It can be used as a helper for finding a certain class of objects. In the application of unknown object detection, we can extend this application to be a warning system. For instance, when some dangerous objects are in a scene, by PCA components analysis of a scene, some abnormal objects will show up in the PCA components as a certain group of objects. In our experiment, top 12 dominated components were analyzed. As a result, the objects such as green pens and rubbers were found among a plant.

The most challenging work in image registration is that two images were taken in different time and from different angles of the view points. We have showed the possibility of the application. In our experiment, a small amount of images have been tested. It is still a challenge when a huge datasets are in a library file and various situations need to be considered. Failure becomes normal. However, if we consider about near future, when multispectral camera becomes available in every smartphone. Synchronized images from video camera and multispectral camera can be matched by a constant displacement. The information overlapping becomes very easy. At that time, multispectral information will be widely used in our daily lives.

SUMMARY

As a summary, this paper proposed an AR technology to enhance video images by superposing the added information from multispectral camera. Multispectral images were taken from red and near infrared spectral bands. These images were explored and analyzed, especially for the NDVI value of the vegetation, object detection from known spectrum, and the PCA components of a scene. These information is stored in a library file as images. A video was taken by a video camera from a smartphone. A similarity measurement image matching method was implemented to check whether an image is matching with the video image. When matching condition meets, the image from the library will be shown up and overlapped on the video. The implementation of the AR system was based on two platforms: Matlab and Android mobile application platforms.

This paper has exhibited the possibility for using remote sensing sensors to enhance a physical scene. EMS beyond human vision offers extra information for our surroundings. And we believe that in near future these information will be beneficial for our lives.

REFERENCES

- Augmented Reality. https://en.wikipedia.org/wiki/Augmented_reality. Last access on 10 February, 2017.
- Dalle Mura, M., Dini, G., & Failli, F. (2016). An integrated environment based on augmented reality and sensing device for manual assembly workstations. *Procedia CIRP*, 41, 340-345.
- Electromagnetic Spectrum. https://en.wikipedia.org/wiki/Electromagnetic_spectrum. Last access on 10 February, 2017.
- Foster, P. J. (2016). *Augmented Reality in the Classroom* (Doctoral dissertation, Loyola Marymount University).
- Human eye. https://en.wikipedia.org/wiki/Human_eye. Last access on 10 February, 2017.
- Hořejší, P. (2015). Augmented reality system for virtual training of parts assembly. *Procedia Engineering*, 100, 699-706.
- NDVI. https://en.wikipedia.org/wiki/Normalized_Difference_Vegetation_Index. Last access on 10 February, 2017.
- Portalés, C., Lerma, J. L., & Navarro, S. (2010). Augmented reality and photogrammetry: A synergy to visualize physical and virtual city environments. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(1), 134-142.
- Wang, Z., Simoncelli, E. P., & Bovik, A. C. (2003, November). Multiscale structural similarity for image quality assessment. In *Signals, Systems and Computers, 2004. Conference Record of the Thirty-Seventh Asilomar Conference on* (Vol. 2, pp. 1398-1402). IEEE.
- Weng, N. G., Bee, O. Y., Yew, L. H., & Hsia, T. E. (2016). An Augmented Reality System for Biology Science Education in Malaysia. *International Journal of Innovative Computing*, 6(2).