kInstitution: University of Huddersfield		
Unit of Assessment: 11 (Computer Science)		
Title of case study: Enabling Social and Commercial Impact from Innovations in Visual Computing		
Period when the underpinning research was undertaken: 2000 - 2019		
Details of staff conducting the underpinning research from the submitting unit:		
Name(s):	**Role(s) (e.g. job title):**	**Period(s) employed by submitting HEI:**
Prof. Zhijie Xu	Professor, Director of Centre for Visual and Immersive Computing (CVIC)	1999 – Present
Dr. Duke Gledhill	Senior Lecturer	2004 – Present
Prof. Paul Scott	Professor	2010 – Present
Dr. Jing Wang	Research Fellow	2013 – 2017
Period when the claimed impact occurred: 2013-2020		
Is this case study continued from a case study submitted in 2014? No		

1. Summary of the impact

Computer vision, visualization and machine learning are state of the art techniques that require enormous computing power. Researchers at the University of Huddersfield have accelerated their practical application by developing techniques that enhanced data processing capability, thus enabling more data sources to be used, the data to be processed more rapidly and more knowledge to be derived.

Key impacts resulting from the underpinning research were as follows:

- The ShenZhen Public Security Department in China installed software that removed the need for human operators for 9,712 CCTV cameras, thus saving £8m.

- The global brand, Phoenix Bicycle was able to update its factories in China to be greener and more efficient, which also opened up new markets, using new computer vision techniques.

- A Chinese supplier of augmented reality-based wearable devices for the visually impaired exploited UoH patents and grew their customer base to 80,000 whilst greatly improving the quality of life for users.

2. Underpinning research

Advances in the overlapping spheres of engineering and computing have been increasingly driven by developments in highly specialised fields such as computer vision (where computers replace the human eye), visualization (computer graphics, animation, and virtual reality), machine learning (where systems improve automatically) and the software and systems strategies associated with them. Practical applications have multiple input streams containing data that must be analysed, combined, and interpreted. A key challenge is ensuring that the data processing burden is not so high that the system cannot be utilized in real-time. Research at the University of Huddersfield (UoH) developed techniques that improved the speed of data processing, which were applied practically in field of visual and immersive computing.

The research described in this case study was carried out at the Centre for Visual and Immersive Computing (CVIC) at UoH. It was led by Professor Zhijie Xu (Director of the CVIC, at UOH since 1999). Other members of the team who aided the impact building included, in surveillance and security (Dr Jing Wang), automation and machine vision (Professor Paul Scott, an associate member of CVIC), and health technology (Dr Duke Gledhill). The work focused on three interlinked strands of expertise. New algorithms and techniques for real-time CCTV analysis; smart factory automation based on machine vision; and improvements in data handling techniques in Augmented Reality (AR) product design used in health products for the visually impaired.

Devising New Algorithms and Techniques for Real-Time CCTV Analysis
Globally, huge amounts of CCTV data are acquired daily by sources such as security cameras (in the public realm), private enterprises such as Tube and Dashcams for insurance firms, and private homes. The efficient processing of this data is now critical in numerous aspects of daily life, not least due to the heightened threat of terrorism and the consequent need for effective surveillance and crowd control.

The research focused on how to use machine learning to enable data from a video camera, to replicate the decisions of a human operator. A foundational study developed a system that used high dimensional information (additional data streams) to identify individual behaviors by constructing compact spatial and temporal signal representations. The techniques learned were further developed and applied to enable real-time intelligent monitoring and early warning of emergencies in large-scale networks monitoring crowds.

The research team devised a representational model of the environment and a corresponding scheme to code the scenario, based on category theory (which enabled disparate data sources to be compared and interpreted) [3.1]. In the model, 'higher' dimensional information such as individual pedestrian 'actions', crowd movement and the nature of the surrounding environment, were used to construct a 'global' model in which a subject's movements in three-dimensional space were inferred by combining multiple 2-D images from CCTV cameras, and then geometrically represented as a whole. This enabled the individuals to be monitored as they moved, including from camera to camera.

When implemented, the model automatically updated and recognized changes in a subject's behavior, motion, and context (environment) by applying a feature map fast vector coding scheme (which ensures features in images can rapidly be identified). The new model used advancements in Artificial Neural Networks (created via so-called deep learning) combined with UoH-developed heterogeneous scalable neural networks (again designed to accelerate data analysis) called the Treble Stream Semantic Network [3.2]. These were used to resolve the long-standing challenge of accurate delicate (subtle) feature (individual item of interest) selection. It was used to process and classify complex crowd scenarios and enabled the recognition of changes in the behavior of an individual relative to their surroundings.

Developing Smart Factory Automation Based on Information Visualization and Machine Vision
Virtual Reality is an important platform for simulating and designing manufacturing systems. Its real-world adoption was slow because complex manufacturing scenarios had to be treated as unique, meaning that each one was individually modelled in the VR environment. This made it particularly expensive to model large scale and complex manufacturing environments, such as assembly lines. The research solved the problem with a novel approach, a top-down virtual environment construction model [3.3], which enabled, for the first time, a modular approach to building complex VR simulations. A software system which separated the "look" (models) and "feel" (GUI and programme functions) of the system, was designed and implemented. A general case was created, which was then applied to multiple manufacturing scenarios, resulting in reduced build time. The model incorporated practical processes that were leading-edge in automated manufacturing at the time e.g., flexible manufacturing concepts such as cells, in a practical manner that led to optimised scheduling, reduced lead time and logistical cost saving.

The research [3.4] explored the whole manufacturing process from design to implementation on the shop floor. It used visualization techniques such as single view and stereo vision to improve the performance of robotics equipment, such as visual alignment for robot guidance (which improved accuracy in part placement), 3D detection (which enabled a robot arm to pick up parts) and parts measurement (used for quality control). Parallel computing hardware acceleration strategies (based on adapting commercially available 'gaming' graphics cards) were devised. These transformed algorithm-acceleration approaches, originally developed for the software arena, to the computer hardware. Consequently, the algorithms ran faster. The faster online processing delivered, meant it was possible to incorporate deep learning capability (i.e., an artificial neural network trained model that runs in real-time).

Deploying Intellectual Property in Augmented Reality (AR) Product Design for Global Visually Impaired Population

Scene recognition establishes the context of the objects in a 'view' (a practical example of which is an ever-changing streetscape). It plays a crucial role in robotic navigation and localization, which is used in assistive technology, especially for the visually impaired. Current scene recognition techniques were prone to error, due to issues with their inability to distinguish individual features and a weakness in template matching (which is used to identify specific elements of ill-defined objects). The research team developed algorithms described in international patents [3.5], which improved i). the integrity of data supplied from multiple sensors (e.g. IR, UV, ultra sound etc.) and provided better signal quality for signal fusion (2016), and ii). an open-source scene recognition programming API [3.6, 5.10], that was made available for programmers to exploit on the IOS and Android platforms. It has transformed practice in the challenging domain of real-time scene recognition by synthesizing data from multiple-input sources in real-time.

New developments included improved differentiation between passable areas and new obstacles (for instance between a wall and a staircase), and recognition of traffic lights and zebra crossings. The approach improved performance versus state-of-the-art benchmarks in the acquisition of close-range depth data. This used laser pulses in a similar fashion to sonar, to create a depth 'map' which was fused with other data sources, such as polarized infra-red (which allowed differentiation between a wall and a glass door). The closest detectable range for the sensors was reduced from 650mm to 165mm with an accuracy of 95%.

3. References to the research

The main outputs in Section 2 have been disseminated through multiple channels, including publishing in top research journals, workshops for industrial and societal stakeholders, invited speeches in prestigious forums and conferences, and leading bids and projects with international partners. References below include high-impact papers, international patents, in-house developed open source toolkits for industry, and Apps for health product users:

1. Wang, J., & Xu, Z. (2016). Spatio-temporal texture modelling for real-time crowd anomaly detection. **Computer Vision and Image Understanding**, 144, 177-187. https://doi.org/10.1016/j.cviu.2015.08.010
2. Xu, Y., Lu, L., Xu, Z., Wang, J., Huang, J., & Lu, J. (2018). Towards Intelligent Crowd Behaviour Understanding through the STFD Descriptor Exploration. **Sensing and Imaging**, 19(17). https://doi.org/10.1007/s11220-018-0201-3
3. Xu, Z., Zhao, Z., and Baines, R. W. (2000) *Constructing virtual environments for manufacturing simulation.* **International Journal of Production Research**, 38 (17). pp. 4171-4191. https://doi.org/10.1080/00207540050205000 [can be supplied on request]
4. Xu, Y., Xu, Z., Jiang, X., and, Scott, P. (2011). Developing a knowledge-based system for complex geometrical product specification (GPS) data manipulation. **Knowledge-Based Systems**, Elsevier. 24(1), https://doi.org/10.1016/j.knosys.2010.05.002
5. **Patent 1**: 201811436184.0 A glass detection device and corresponding methodology based on fusion of RGB-D and ultrasonic sensors; and, **Patent 2**: 201811436208.2 A scene representation device and methodology based on semantic stixel. https://bit.ly/3qUERYQ [Patent applications available on request]
6. **Development API and Apps** https://a.app.qq.com/o/simple.jsp?pkgname=cn.krvision.litekrnavi; and, **Software API:**

OpenMPR - the open-sourced software for place recognition using multiple descriptors derived from multi-modal images: https://github.com/chengricky/OpenMultiPR

4. Details of the impact

The research findings led to commercial impacts in three organisations in China. The impacts can be summarized under three headings: 1. New forensic imaging technologies for law enforcement; 2. Enabling automation in a bicycle factory; 3. Commercialization of wearable devices for visually impaired individuals.

New Forensic Imaging Technologies for Law Enforcement

CCTV is a pervasive tool for monitoring crowd behaviour and has had proven success in enabling the detection of suspicious individuals, such as terrorists and alerting the authorities to risks to individuals in large crowds. Each camera, however, needs to be monitored by a human operator, which introduces a non-scalable cost factor.

Forensic imaging scientists at the Shaanxi Electronic Information Scene Investigation (EISI) in China collaborated with the UoH CVIC since 2010. One project [5.1] resulted in the creation of a real-time crowd behaviour analysis capability [3.1] (2016). It was trialled as a live system in the southern Chinese city of ShenZhen and significantly reduced the number of people required for continuous (24-hour) real-time monitoring. It also improved the emergency response speed for public emergencies and reduced loss of life and damage to property [5.3].

The technology [3.2] and IPs [5.2] were adopted by policing and public security bureaus in eight provinces in China [5.4]. In one example, the system has been installed in more than 640 sites (such as streets, railway stations and parks) in Shenzhen City, Guangdong Province, with 9,712 video cameras in the network. Since 2016 more than 22,000 abnormal emergencies have been identified. Multiple police, traffic control and crime prevention agencies provided positive feedback on the pilot system, which delivered an estimated cost-saving to the public purse of 70 million RMB dollar (£8m) since 2016. Significantly, it is estimated that in the Sichuan Province alone, the technology saved over 3,000 security monitoring personnel. Based on the average cost of a police officer per year, the technology has reduced costs by up to 240 million RMB dollar (£27.5m) annually for the provincial public security and police departments in Sichuan. The Police Superintendent stated "… the crowd monitoring technology significantly reduced the cost of police resources including time, workload, and public grant" [5.3].

Enabling Automation in a Bicycle Factory

The China-based Shanghai Phoenix Company Ltd. produced 4.89 million bicycles in 2019 and exported to more than 50 countries. In 2000 the manufacturing practices of the company were similar to those seen in Europe in the 1920s. Since then, the company has collaborated with various global research institutions to accelerate the modernisation of its production lines. The CVIC at UoH provided research and consultancy-based support since 2010, focused on the design and development of an efficient shop floor by utilizing virtual manufacturing-based planning and simulation [3.3]. This enabled Phoenix to modernize its production operations to match current global best practice in just a decade, becoming more energy efficient and cost effective [5.5].

The CEO of Phoenix stated "Prof. Xu's expertise in industrial design and automation has greatly reduced the lead time to modernise and operate complex manufacturing and production environments. Substantial cost saving from labour and energy of up to 15 percent have been recorded in the past 3 years. The increased production capacity has also enabled the strategic redirection of the product type towards more green lines. For example, the partnership with the Shared Bicycle scheme (OfO Ltd.) exported Phoenix products to countries including France, United States and United Kingdom" [5.6].

Visual guidance systems were jointly developed with the sector leading Second2None Machine Vision Systems Co. Ltd [3.4]. These included the 'Long-Rui' series of smart vision sensors, which were applied to a series of vision appliances used for robot guidance, 3-D detection, visual alignment, and welding control in two of Phoenix's plants. The company has been able to sell the new product to Epsom and Yamaha. It has grown for five years in a row and was named a Star Company in Shenzhen City [5.7].

Commercialization of Wearable Devices for Visually Impaired Individuals
The research facilitated ShiKe (KrVision) Ltd to commercialize wearables and headsets for the visually impaired in 2016. It exploited patented technologies [3.5] [5.8] and the open source API [3.6]. The product featured cutting-edge functionality in scene localization, video augmentation and sign and text recognition, plus the capability to render the information synthesised from the data in audio form. It could remember a regular route followed by the wearer, which was used as a template that enabled greatly improved identification of new obstacles. In addition, the device expanded the 'visual' range of a totally blind person from the 'point in space' defined by a 'white stick', to a three-dimensional hemisphere that radiated from the front face of the wearable device. The founder of the company stated, "… the collaboration with the CVIC at UK has greatly accelerated the release of our new AR kit" for the visually impaired community [5.9]. He added "The improved functions and new features on audio coding and feedback, semantic segmentation, street water pool and sinkhole identification have greatly benefitted the users and could not have been implemented so successfully without the collaboration with our international partners". As a result of the joint effort, the company has grown from an SME in 2014 into a medium sized enterprise with a customer base of over 80,000 users across world [5.9].

A visually impaired user of the wearable device said "My life has been utterly transformed by my new headset. I can now move about more easily outside my home, confident that I will have a better warning of any obstacles I need to avoid" [5.8, 5.10].

5. Sources to corroborate the impact

5.1. Shenzhen Project Report. GJHZ20160301164521358. https://bit.ly/2Mq7Owl
5.2. 6 International Patents on live indoor-outdoor individual and crowd abnormality detection https://bit.ly/3qUrDLL
5.3. 2020 Sichuan Science and Technology Progress Award. Project No. 55 https://eur02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fhelios.hud.ac.uk%2Fsengdg2%2F2-provincial-award-certificate.pdf&data=04%7C01%7Cz.xu%40hud.ac.uk%7Ce46f21f6e2e04f4d745b08d8bd521373%7Cb52e9fda06914585bdfc5ccae1ce1890%7C0%7C0%7C637467507495286075%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C1000&sdata=iST54MwT2ZmW8RJK%2B5U%2FGTI41RyNQCk2LBMjd5fPcZc%3D&reserved=0
5.4. Chief Scientist, National Key Laboratory of Electronic Information Scene Investigation (EISI), Endorsement Letter
5.5. News Report: https://www.instrumentation.co.uk/switch-to-new-smart-factory-technology-ticks-greener-manufacturing-box-for-bicycle-production-giant/
5.6. Chief Executive Officer, Shanghai Phoenix Enterprise (Group) Co.Ltd., Endorsement Letter
5.7. Deputy Chairman of the China Machine Vision Union, Managing Director, Second2None Machine Vision Systems Co. Ltd., Endorsement Letter.
5.8. Collaboration Evaluation Report, KrVision https://bit.ly/39fmdVO
5.9. Founder and Technical Director, ShiKe (Kr Vision) Co., Ltd. Endorsement Letter
5.10. Products, Awards, and User Feedback, ShiKe (Kr Vision) Co., http://www.krvision.cn/offical/page/assist.html