IRSTI 50.01.85

S. Kalenova[1], Zh. Duisebekov[2]
[1,2]Suleyman Demirel University, Kaskelen, Kazakhstan

NOVEL HUMAN-COMPUTER INTERACTION THROUGH A PRISM OF COMPUTER VISION

Abstract. For the past two decades, the innovative developments in Computer Science have made the crucial beneficial influence on Human – Computer Interaction (HCI) applications. Recent advanced HCI techniques are essential elements to enhance current computer interfaces and facilitate the fundamental user-friendly interaction. It is now becoming possible to communicate with computers by using hand gesture, pose recognition, eye tracking, emotion recognition and several other developed progressive technologies. This paper investigates the direct contribution of Computer Vision (CV) applications to improving various modes of human-interaction, specialized in different prominent industries, where cumulative list of use cases includes sport, healthcare, agriculture, transportation, retail and manufacturing fields. The current work also explores the future forecast of CV applications for the next decade and identifies possible scenarios of research directions.

Keywords: Human – Computer Interaction (HCI), Computer Vision (CV), hand gesture, pose recognition, eye tracking, emotion recognition.

***

Аңдатпа. Соңғы онжылдықта компьютерлік ғылымдар саласындағы инновациялық дамулар адам мен компьютердің өзара әрекеттесуі (HCI) қосымшаларына шешуші тиімді әсер етті. Жуырдағы жетілдірілген HCI әдістері қазіргі компьютер интерфейстерін жақсартудың және пайдаланушылармен ынғайлы қарым-қатынасты жеңілдетудің маңызды элементтері болып табылады. Қазіргі кезде компьютермен қолмен қимылдау, позаны тану, көзді қадағалау, эмоцияны тану және басқа да бірнеше дамыған прогрессивті технологияларды қолдану арқылы байланыс орнатуға болады. Бұл мақалада спорт, денсаулық сақтау, ауылшаруашылығы, көлік, бөлшек сауда және өндіріс сияқты әр түрлі өнеркәсіптік салалардағы адамдармен өзара әрекеттесудің әртүрлі тәсілдерін жақсартуға компьютерлік көру (CV) қосымшаларының тікелей үлесі зерттелген. Ағымдағы жұмыс сонымен қатар алдағы онжылдыққа түйіндеме қосымшаларының болашақ болжамын және болашақ зерттеу бағыттарының ықтимал сценарийлерін жобалайды.

Түйін сөздер: адам мен компьютердің өзара әрекеттесуі (HCI), компьютерлік көру (CV), қол қимылы, позаны тану, көзді қадағалау, эмоцияны тану.

***

Аннотация. За последние два десятилетия инновационные разработки в области компьютерных наук оказали решающее благотворное влияние на приложения взаимодействия человека с компьютером (HCI). Последние передовые технологии HCI являются важными элементами для улучшения существующих компьютерных интерфейсов и упрощения фундаментального удобного взаимодействия. На данный момент, представляется возможным общаться с компьютерами посредством распознавания жестов рук, определения позы, отслеживания взгляда, распознавания эмоций и ряда других прогрессивных технологий. В этой статье исследуется прямой вклад приложений компьютерного зрения (CV) в улучшение различных способов взаимодействия с людьми, специализирующихся в различных индустриальных отраслях, таких как спорт, здравоохранение, сельское хозяйство, транспорт, розничная торговля и производство. Текущая работа также исследует будущий прогноз приложений CV на предстоящее десятилетие и проектирует возможные сценарии направлений будущих исследований.

Ключевые слова: взаимодействие человека с компьютером (HCI), компьютерное зрение (CV), жесты рук, определения позы, отслеживания взгляда, распознавания эмоций.

Introduction

To start with, in July of 1945, the American engineer Vannevar Bush wrote an article for the Atlantic Monthly, called "As We May Think", where he outlined the use of computers beyond calculation and share the true vision for information technologies that should transform people's intellectual abilities for raising their creative activity [1]. He predicted the use of digital devices as a fundamental tool for multimedia processing. Nowadays, computers are human-centered mostly via inputs, e.g., keyboard and mouse, however there are few of them, which are able to handle information from audio, visual or haptic inputs [2]. Since visual perception is the best instrument in communication between organisms and environment, it becomes obvious, that Computer Vision (CV) might bring the people and computers together and help to build a closer connection. Current interaction based on body language and gesture recognition techniques.

As for basic definition, CV is a branch of computer science that related to Artificial Intelligence (AI) and uses machine and Deep Learning to help computers to "see" and process surround data. The recent technological findings in CV computing have changed the Human – Computer Interaction (HCI) astonishingly. The massive impact of CV to HCI applies to issues where the situation needs a human's eye to monitor activities. Therefore, the term "vision" brilliantly describes the core mission of CV that more preferentially works than simple human factor.

This paper reviews the latest progressive applications of computer vision for HCI that used in various industrial areas in 2021. The fast-growing CV use cases are notably discovered in the next fields:

1. Sport (Player/Multi-Player Pose Tracking, Markerless Motion Capture, Objective Athlete Performance Assessment, Stroke Recognition, Ball Tracking etc.);

2. Healthcare (COVID-19 diagnosis, Cancer Detection, Cell Classification, Mask Detection, Disease Progression Score etc.);

3. Agriculture (Flowering Detection, Plantation monitoring, Agricultural Product Quality Testing, Animal Monitoring, Farm Automation etc.);

4. Transportation (Pedestrian Detection, Automated License Plate Recognition, Traffic Flow Analysis, Vehicle Classification, Driver Attentiveness Detection etc.); 5. Retail and Manufacturing (Customer Tracking, People Counting, Theft Detection, Theft Detection, Waiting Time Analytics etc.) [3].

Computer Vision in Sport

The novelty of combination of Artificial Experience and Computer Vision algorithms is that now it became possible to recognize human body language and pose movement patterns through both monocular, defined as single-camera footage, and multi-view, which footage of multiple cameras, sport datasets, that usually collected from archive videos or live broadcasts. Real-time single and multi-person 2D or 3D pose estimation algorithms opens up a new opportunity for machines to intuitively understand players performance analysis, motion capture without manual annotation the body segments in each video frame. The popular uses cases are swimming analysis and prediction of basketball free throw shooting. In the first case Convolutional Neural Networks (CNN) are used to automatically extract the required pose data and detect the swimming style of athlete. The second application builds a posture analysis model using OpenPose skeletal recognition data, where applied a simple logistic regression model that predicts the shooting probability of the basketball free throw with skeleton posture data as explanatory variables and the fact whether the ball enters the basket or not as a binary target variable [4].

Computer Vision in Healthcare

Since occurred the most dramatic issue of 2019-2020 years, called COVID-19, CV engineers vigorously worked at development of global coronavirus disease control and monitoring software. Multiple deep learning computer vision models exist for x-ray based COVID-19 diagnosis. The most popular one for the detection of COVID-19 cases with digital chest x-ray radiography (CXR) images is named COVID – Net and was developed by Darwin AI, Canada [3]. In addition, Machine Learning (ML) with CV intersects in medical industries aimed for breast and skin cancer detection. With help of image detection, scientist identify cancerous and non-cancerous results, and

diagnose data from magnetic resonance imaging (MRI) separately for malignant or benign.

Computer Vision in Agriculture

Smart agriculture can be used for different purposes, starting with plantation monitoring, insect and animal detection till overall farm automation (Fig. 1). For instance, timely insect detection is the great contribution in pest control. Traditional manual identification and counting of flying insects is inefficient and labor-intensive. Vision-based systems allow the counting and recognition of flying insects (based on You Only Look Once (YOLO) object detection and classification). As for animal monitoring, that application helps in remote monitoring and recognize various diseases safety at a distance.

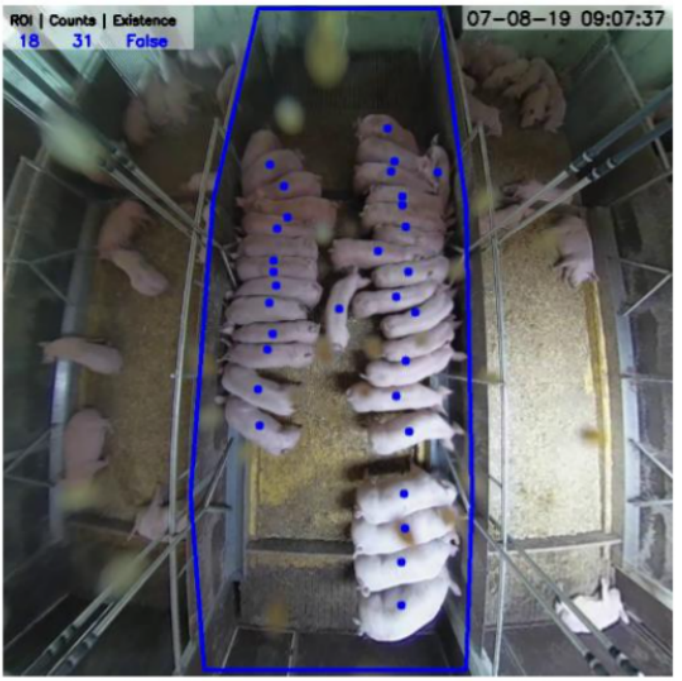

Figure 1. Animal monitoring

Computer Vision in Transportation

Today automated vehicle license plate recognition and pedestrian detection are basic and familiar CV use cases for convenient HCI. Numerous modern public safety systems are integrated with plate recognition algorithms from input images and videos. One of the most popular libraries for vehicle registration plates is OpenALPR, that based on character recognition on images or video feeds (Fig. 2). As for pedestrian detection, there are many input censors for their recognition, such as CCTV / IP cameras, thermal imaging devices and simple face recognition welcome door apparatus. Pedestrian detection algorithms can be based on infrared signatures, shape features,

gradient features, machine learning, or motion features. Thus, different automated road items monitoring systems can be observed as complex intervention to human safety issues by producing quick transparent computer interaction.

Figure 2. Vehicle license plate recognition

Computer Vision in Retail and Manufacturing

HCI has also entered the world of Retail and Manufacturing in the forms of customer tracking (Fig. 3), people counting, social distancing, and workspace productivity analytics. Comes back to recent situation regarding COVID-19, when only a limited number of people were allowed in shops and market at the same time, moreover social distancing aspect. Due to these, many companies established camera detectors, where camera tracked employee or customer movement and uses depth sensors to assess the distance between them. Depending on their position, the system drawn a red or green circle around the person.

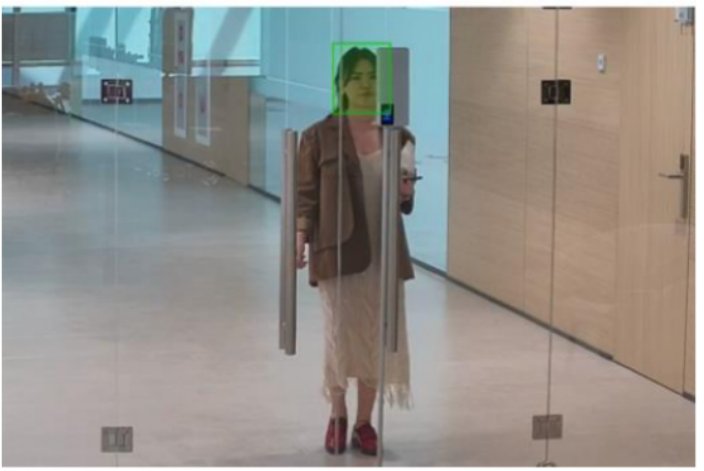

Figure 3. Face recognition

The Future of HCI

The most intriguing prospective computer interface is direct brain-gadget communication. This will require accurate mapping of neuronal signals in our brains so that we can communicate with and control external devices - just by thinking about the command. In a reality, it is complicated task to accomplish, because human brain more complex than we may assume, also to reach brain sensors requires a lot of invasive insertion of electrodes. However, current deep learning algorithms are already ready to train challenging models and provide a proper feedback from computer AI brain, therefore it is clear that novel HCI subjects will mainly work on improving CV interaction channels.

Conclusion

To conclude with, latest Human – Computer Interaction innovations have made many human-centered solutions, especially via Computer Vision support. The scope of CV impacts covers almost all daily life issues in sport, healthcare, agriculture, transportation, retail and manufacturing. The main objectives of this paper were: 1) to identify the direct CV intervention to HCI techniques; 2) propose the possible future HCI interfaces and interaction channels. In sum, transparent truth is that current progressive CV applications totally made easier the HCI communication. Appropriately, literally CV is the most powerful tool that acts as a bridge to a future "mental" human – computer communication.

References

1 Bolt, R.A. Conversing with computers. Technology Review, 88 (2), pp. 35-43.

2 Homero, V. Rios. Human-computer interaction through computer vision. Xalapa, Veracruz 91000, Mexico, Universidad Veracruzana, pp. 1-2.

3 Viso.ai, 56 Most Popular Computer Vision Applications in 2021, 9 March 2021, accessed 15 May 2021, URL: https://viso.ai/applications/computer-vision-applications/ >.

4 Masato, N., Yoshihiko, T., Hisashi, H., Hideki, M. Prediction of Basketball Free Throw Shooting by OpenPose, *School of Industrial Technology, Advanced Institute of Industrial Technology* 1-1040, Higashiooi, Shinagawa-Ku,Tokyo, 140-0011, Japan, pp. 2-5.