Agents of M.A.S.K.: Mobile Analytics from Situated Knowledge

With increased digital sensor platforms and infrastructures emerging, we are seeing a massive increase in data being gathered from objects, processes, and spaces. For example, social media interactions, data collected from IoT devices, financial transactions, transportation-related data, governmental records, and scientific datasets all make substantial contributions to the expanding pool of data sources.

In order to provide personalized, effective, timely, and responsive services, this growing volume of data needs to be fielded, sorted, and rendered accessible anytime and anywhere. This will allow users to interact with adaptive physical objects, spaces and their associated data [1]. Recently, immersive [2] and situated analytics [1] have been introduced to expand the visual analytical space using virtual and augmented reality technologies. These technologies build analytics on the user’s capacity to move in the physical world, between different environments and scenarios. However, a critical challenge arises when deploying analytics in linked query sessions that are connected to different virtual or physical worlds. In these cases, users need to switch between dynamic display spaces, various interaction modalities, and fragmentation of analytics sessions.

The emerging field of mobile analytics advocates the immediacy of adapted situated analytics based on knowledge of the surrounding environment. To showcase the challenges (C1—C6) and research questions (RQ1—RQ2), we present an illustrated use case for immersive and situated analytics, highlighting the technology limitations that prevent the deployment of analytics in different environments, Examples we illustrate include: i) illumination, ii) clutter, iii) static vs. dynamics of multiple objects in motion, iv) close-up views, v) wide-open views, and vi) first-person motion. These aspects make it challenging to deploy arbitrary mixed reality techniques in real-life connected scenarios.

Our illustration which takes the form of a speculative design scenario presents a vision that merges situated analytics with artificial intelligence and behavior analysis to empower analytics on-the-go. Mobile analytics implies an expectation of immediate availability, while augmented reality enables existence within the physical environment. We draw the techniques and methods needed to offer a fully immersive experience, with artificial intelligence, generating a cohesive personalized presentation for mobile analytics.
THE STORY OF A MASK

It all began when an important piece was stolen from a museum. Detectives must solve this case quickly and accurately, as any incorrect analysis path can lead to losing the piece. This is a typical visual analytics case, where detectives need to analyze different datasets using diverse models (artificial intelligence (AI) engines), supported by visual representations [3].

The team gathered around the "Revealing Pool", a table-like device for real-time holograms, much like a workbench or showcase. They revealed the digital twin of the missing piece, an almost physical but ghastly object. Instantly, the space around object illuminated with crowd data, illustrating patterns of motion at different time intervals. The pool's surface seems liquid, but it is tangible, made of innumerable particles that raise/sink or adopt color according to the data.

The case was assigned to the Digital Detective Intelligence Agency (DDIA). The team gathered in the main control room, which is equipped with the most advanced immersive analytics tools.
The interface allowed the Gwish to navigate through time forwards and backward. With the advantage of infinite locomotion, the Gwish was able to explore the museum like a regular visitor at crime time!

The Gwish interleaved activities in connection with the analysts’ queries. The suspected visitors highlighted on a large screen, while desktop analytics showed facial expression heatmaps in conjunction with the locomotion patterns. Concurrently, trajectories and other demographic statistics of the subjects appeared in an immersive multiuser tiled display. Analysts went through the spatiotemporal data combined with demographics, pulling the visitors’ data, to and routes.

The head of DDIA was monitoring various feeds in an old-fashioned multi-display analytics setup. With the list of suspects narrowed down, a list of locations, objects, and people emerged. Just as the media was preparing to make the announcement, the decision was made to bring in a field agent, and Agent "A" was the best option!
Not to make "A" isolated from control room of DDIA, he got the **MASK**. The **MASK** was built for **AR** mobile analytics, considering physical world challenges, such as:

- Tracking objects of interest amongst multiple moving targets,
- multimodal clutter of environment,
- depth perception of augmentations,
- visual coherence, e.g., color and illumination,
- the graphical widget relationship with physical objects,
- occlusion of real-world with augmentations, and
- attention and distraction to augmentations,

Agent "A" wore and calibrated the mask. The mask has eye tracking and connects to haptic gloves. The eye tracking controls the presented content based on "A"’s performance and uses for gaze interaction. The haptic glove provides vibration, thermal, and kinesthetic feedback, and tracks "A"’s hands. As, the actuators and sensors drain the gloves’ energy, the actuators were disabled and can be activated manually if needed.

The mask supports data analyses in immersive environments: virtual reality (**immersive analytics**) [3] and augmented reality (**situated analytics**) [1]. With a hand gesture, "A" initiated the mission on the mask. The accurate tracking supported by the gloves enables "A" to carry out the interactions in secrecy. DDIA asked "A" to collect data on two suspicious visitors, from the field, which could not be collected digitally.

Once, "A" went out of the research lab, the mask invoked the outdoor tracking mode, presenting the task’s information on augmented 2D canvas. The augmented window was chosen to reduce the tracking challenges as "A" will be walking. Moreover, drop tracking can help in reducing the processing consumption.
"A" used a mid-air interaction with the presented content. However, to keep the interaction inconspicuous, "A" can reach far sections with gaze and glove interaction. The mask supported "A" by hand avatar, enhancing the near-field interaction.

Audio streaming from the DDIA control room uses spatial sound localization. However, things were busy at DDIA Head-quarter (HQ), which produced a cluttered audio environment, disrupting the automatic sound localisation. "A" opened and control the audio manually, and he kept manually switching between the audio sources to communicate with different teams in the control room.

"A" was guided by the mask through the city with virtual cues, while collaborating with DDIA Headquarter. However, when "A" arrived at a busy street the navigation arrows started to clutter with the surrounding, occluding the physical scene. Moreover, multiple objects moving in the dynamic environment and the self-movement of "A" started to confuse the alignment and contrast of the situated visualisation. The mask used a context-aware technique to update to appearance and location of the annotations, however, as "A" and the scene were rapidly changing, the adaption frequency was very high.
"A" was tasked to check a trace of phosphorescent paint on two suspicious individuals ("S1", "S2") who had been selected based on the analysis session in the control room. "S1" and "S2" were narrowed down by the analysts. "A" started his way to the first location. "A" found "S1" while he was having a coffee. "A" scanned the phosphorescent trace on "S1" using the mask’s built-in Ultraviolet (UV) camera, but he did not find any trace.

However, when "A" reached "S2" he found a trace of the paint on the suspect’s hand and clothing. The computer vision algorithm in the mask recognized "S2" and enabled the analysis mode. In this mode "A" could filter the data in the DDIA HQ remotely. By selecting "S2" in the physical world, as major suspect. This situated analytics brushing action updated the data in the DDIA HQ. A massive cascade of smart filters took place at all stations.

The museum video footage was trimmed to include only the shots with "S2". The DDIA agent at the "Revealing Pool" marked a person ("S3") that was around "S2" most of the time. Selecting "S3" filtered the spatiotemporal data in front of the analysts, and the trajectory data on the wall display.

Real-time videos of the traffic light CCTV cameras were processed, which found "S3" in a traffic light. The team processed the predicted trajectories between "S3" and "S2", estimating "S3"s car path, which properly contains the stolen piece.
The vehicle type was recognized in CCTV frames. The team fed the vehicle's parameters and physics to a 3D simulator to estimate the load and position of the stolen piece in the vehicle. Subsequently, AI algorithms were running analyses using mesh deformation models to figure out how to stop the car with the least impact on the car and the stolen piece inside.

The predictive simulation calculated an optimal hitting point, angle, and car speed. The car must be shot in the back right tire, at angle 37° at car speed 52 Km/h. The mission of pulling the car over was assigned to "A" as he was the one nearest to the car area. However, "A" needs to be trained on the correct shooting angle, as any mistake may cost them the stolen bust! "A" received a call to join a VR training session through the mask. A DDIA expert will teach him how to stop the car safely. "A" accepted the training and started the VR session. The mask enabled "A" to visualise and interact with the car’s digital twin.

The VR training used hand gestures and haptics, to increase engagement. While aiming and shooting the car tire in the virtual training, isolated from the real scene, "A" accidentally bumped into a walking man on the street. "A" finished the training, and the mission changed from exploratory analysis to operation analysis.

"A" was given two options to chase the suspicious car, either by car or motorbike. The mask presented the impact of each option on time, however, the massive, presented information cluttered the physical world. Such clutter representation usually uses interactive visualization, clustering the data into an overview and details on-demand view. However, that is challenging as "A" is walking! Especially for dynamic and cluttered environment.
Situation Description Challenges’ Topic

..."A" opened and control the audio manually, and he kept manually switching between the audio sources to communicate with different teams in the control room...

C1: Manual Adaptation: User required to adapt interactive dashboards’ settings, in his limited physical interaction space.

Manual controls, dexterity, and out of reach controllers.

... The mask used a context-aware technique to update to appearance and location of the annotations, however, as "A" and the scene were rapidly changing, the adaption frequency was very high...

C2: Occlusion: The presented information was occluding physical world’s objects, which were moving faster than the system could adapt.

Information overload, adaptation speed, latency, and dynamic canvases.

..."A" was tasked to check a trace of phosphorescent paint on two suspicious individuals ("S1", "S2") who had been selected based on the analysis session in the control room ...

C3: Disconnected sessions: While the user scanned the physical world, the streamed visualization was disabled, helping him to focus on the surroundings. However, this action disconnected him temporary from the remote collaboration.

Limited view space, and asynchronous remote collaboration.

... The team processed the predicted trajectories between "S3" and "S2", estimating "S3"’s car path, which properly contains the stolen piece...

C4: View representation: All analysts, including those with different backgrounds such as Egyptologists and detectives, were exposed to the same visual representation, which influenced their behavior during the collaborative analytics session.

Adaptive visual representation, and behavior awareness.

..."A" accidentally bumped into a walking man on the street...

C5: Isolated view: The user was isolated in VR mode with no awareness of the dynamic physical environment surrounding him.

Peripheral awareness, and view’s dynamic transition

... the massive, presented information cluttered the physical world...

C6: Clutter view: The flood of information added to the physical surroundings created a cluttered view that distracted the user.

Information overload, density of presentation, and augmentation layout.

"A"’s trip was easier using the mask, however, some of the techniques were not effective in-the-wild! So "A" goes to a rendezvous point, where an update for the mask’s software was handed over. This update enhances the AI situated mobile analytics using dynamic analytical interface based on user behaviour, situation, and surrounding knowledge, including physical world’s features, and displaying mode.
Using the new update, the visualisation blended into the physical world, reducing the scene clutter. The visualisation aligned and scaled based on the task and allocated to the scene based on its expected purpose. For instance, the important data presented as fixed-location abstract data, while the complex graphs presented as situated visualisation, which scaled to enhance on the visual perception.

When "A" arrived near the suspicious car zone. The mask switched to the secret interaction mode, changing from hand gestures to embodied interaction, keeping "A" inconspicuous. "A" analysed two given choices which were calculated remotely at the control room. "A" could select, filter, and analyse data nodes. "A" can “select and “filter” in real-time using foot interaction. "A" stepped on "option one", viewing the data’s details on demand (DoD). The detailed view appeared as canvases blended in the surrounding, and the overview was presented as blended color code. "A" considered the car option, which was faster and more secure for car chasing. However, based on his field observation, presented traffic map and VR training, it would be difficult to arrive obtain the required shooting angle using the car. Therefore, "A" chose the motorbike.
"A" selected the transportation option by standing in the "bike zone" in the floor's interaction area. The mask reverted from analysis mode to execution mode. In execution mode, the mask allocated the canvases on the side buildings, reducing clutter.

The abstract data was presented on the floor using a mediated reality technique, dimming the least important parts of the physical world. "A" followed the augmented cues, while the mask kept measuring "A"'s bio-signals. Sensor's data were used to adapt the task sequences. The built-in eye tracking used to adapt the visualisation based on "A" gaze movement. "A" signal measurement and gaze interaction were sent to DDIA control room, to measure his performance accurately, assessing him throughout the task.

"A" selected the transportation option by standing in the "bike zone" in the floor's interaction area. The mask reverted from analysis mode to execution mode. In execution mode, the mask allocated the canvases on the side buildings, reducing clutter.

Once "A" reached the motor bike, the mask switched to the rapid interactive visualisation mode. The mask enabled only selection option, which can be controlled by eye gaze, and glove click, with blended information representation. "A" can select the object infront of him, to explore the attached information.

"A"'s bike was capturing the car's speed (using build-in sensors), and the car's orientation (using the mask's camera). The captured data was streamed to DDIA to be processed and calculate the vehicle load's status using the physics engine, fast GPU, and distributed rendering.
Based on the physics engine calculation, the mask augmented for "A" the optimal shooting trajectory, as situated lines to his hand to reduce the rendering cost.

In the meanwhile, one of the DDIA team was manipulating the city traffic to slow down the suspicious car. "A" was streaming the reality's live scene using the mask’s camera, combined with the eye tracking.
The presented storyline gave a comprehensive look at the combination of mixed reality and visual analytics for mobile analytics, emphasizing the immediacy requirements for information on-the-go. We advocate the need for a fully immersive experience that considers both the real-world situation and the abstract information, with artificial intelligence generating a cohesive personalized presentation of real and virtual. Various paradigms for the access and analysis of digital information have been put forth: visualisation [4], visual analytics [3], immersive data representations [5], immersive analytics [2], situated analytics [1] and embedded data representations [6]. Mobile analytics implies an on-the-go expectation of immediate availability, while augmented reality enables existence within the physical environment. We draw from the body of literature the techniques and methods needed to offer a fully immersive experience, with artificial intelligence, generating a cohesive personalized presentation of real and virtual.

The storyline highlights the existing visualisation, interaction, and analysis for mobile analytics, and its potential research directions (R), which can be summarized as follow:

**RQ1. Adaptive UI blended interface based on user behaviour**—by blending the augmented visualisation into the physical environment, enabling system to manipulate the view based on situated knowledge, such as object tracking, depth, and clutter factor.

**RQ2. Adaptive embodied interaction**—by dynamically changes its interaction mode based on the user’s behavior and surrounding knowledge.

**RQ3. Blended visualisation for perception enhancement**—by employing scene manipulation techniques to control users’ attention. These models can be automatically adapted based on situated knowledge, including factors like clutter percentage, depth, illumination, and more.

**RQ4. Visual cues for dynamic situation augmentation**—by utilizing a predefined static models to adapt the situated visualisation for fast animated, leveraging the behavioural knowledge of users’ status.

**RQ5. Task-driven analytic interface**—by adapting the situated visualisation and adjust the interaction tool based on predefine task goals.

**RQ6. Collaborative mobile analytics**—by dynamically adapt the visual analytic session for each user independently based on users’ behavioural knowledge.

In this feature article we have introduced, reviewed, and illustrated various parameters affecting mobile analytics in-the-wild. Through our illustrative scenario, we presented the challenges and requirements for deploying mobile analytics in real-life scenarios. We envision the impact of adapting visual (situated and immersive) analytics when considering “Situated Knowledge” and “Artificial Intelligence (AI) user interface”. Situated Knowledge includes physical environment’s parameters, analytics’ tasks, and user’s behavioural data. AI engines uses the collected parameters for generative models, calculating the augmentation settings and control models.