Based on the VR display project of the museum, this paper elaborates the temporal restoration, spatial restoration, and immersive experience of virtual reality technology and discusses the basic research of virtual reality technology in the museum, the display of the platform, and the prospect of cultural communication. By studying the working principle of somatosensory interaction technology, the development status of the somatosensory interaction device Kinect, and the programming algorithm, this paper explores its development trend and application prospect in the field of interactive exhibit design and production in science and technology museums. A prototype system of human-machine interaction is realized in the exhibit design, using Unity3D as the engine to build the interaction platform, and using webcam and other interactive devices to realize somatosensory interaction, improve the design scheme of interactive science exhibits, and practice diversified and multisensory innovative exploration of science forms.

1. Introduction

Most museums still maintain the traditional form of object orientation, displaying exhibits in cabinets with basic labels and then ending the educational mission [1, 2]. Although countless “high level” “fine art” collections are on display, the value of the collections is hardly understood, and for most visitors, it is just a visit to the museum. At a time when traditional means of display are struggling to meet the educational needs of museums, the rapid development of new media technologies has played an important role in shaping and disseminating popular culture. It is for this reason that many new media technologies have been introduced into museum display design and have become part of the display language of museums, making a breakthrough in the educational model of modern museums [3, 4]. The advantages of new media technologies applied to museum education are mainly as follows.

Many museum exhibitions due to the relatively limited collections and their historical materials, even if the theme is original and intentional high, but still due to the limitations of traditional display methods, cannot show the wonderful story behind the collection; visitors to the what visitors want to see most in museums are the precious prototypes of cultural relics, especially for some of the town treasures, and they often come here by name [5]. When viewing the exhibits, the audience is more willing to understand the history and artistry of the exhibits themselves. This information is usually not limited to the exhibits themselves, and it is difficult to express it through traditional display methods such as static text and pictures; those historical relics and cultural landscapes that have been destroyed by nature or man-made long ago, it is difficult to use conventional methods. It is difficult to reproduce the scenes [6, 7].

As mentioned earlier, museum learning is a free choice learning, the educational environment of the museum is not
the educational environment of the museum is nonmandatory, and visitors have a strong sense of autonomy in their behavior in the museum. Museum visitors are at many different levels. Museum visitors are of many different levels, reflecting multi-level, multifaceted, and diverse information needs, but the traditional display of exhibits, mostly with in this single exhibition mode, the information transmission lacks specificity, and it is difficult for the museum to respond to the different aesthetic needs of visitors. It is difficult for museums to differentiate their displays to meet the different aesthetic interests and knowledge needs of visitors [8]. With the introduction of new media technology, visitors can use their own mobile devices to search for information about the collections they want to know, and they can also use some of the Jewish Museum of New York, on the other hand, has a number of electronic touchscreens in its exhibition areas [9]. The Jewish Museum in New York has installed a multimedia device in front of each gallery, which automatically introduces the contents and features of the exhibition to visitors entering the gallery. The museum has installed a multimedia installation in front of each gallery, which automatically introduces the contents and features of the exhibition to visitors entering the gallery, so that they can freely choose the content of their visit according to their preferences and different needs. The visitors can freely choose the contents of the exhibition according to their preferences and different needs. Through the new media technology, visitors can independently exhibit and select the exhibits they need to know this way; visitors can choose the exhibits they need to know and learn more about them, thus improving the efficiency of their visit. The diversity of new media technology and the richness of the information it carries have obvious advantages in meeting the individual needs of modern visitors and stimulating their participation and enthusiasm. It also stimulates the participation and enthusiasm of visitors [10].

As one of the most popular human-computer interaction methods in the 21st century, somatosensory interaction has been developing rapidly in recent years. Somatosensory interaction usually requires a series of technical support such as motion tracking, gesture recognition, motion capture, and facial expression recognition. Compared with other means of interaction, the current body-sensory interaction technology has been greatly improved both in terms of hardware and software. Interaction devices are becoming smaller and more portable, and easier to use. At the same time, no direct contact is required in the interaction process, which greatly reduces the constraints on the user and improves the immersion of human-computer interaction, making the interaction process more natural. Augmented reality (AR) is a technology in which computers project virtual objects and scenes into real scenes and interact with the surrounding environment to achieve “augmentation” of reality, which emphasizes the combination of virtual and reality, i.e., “virtual-real.” In recent years, AR and Kinect have become a hot topic for many researchers, and the combination of Kinect and augmented reality technology for the design of interactive exhibits in science and technology museums is a hot topic for research [11, 12].

Currently, new media technologies have a variety of manifestations in museum displays, ranging from technically more mature individual examples, but they do not clearly reflect the overall picture of new media technology integration in museums. The key to the application of new media technology is the technology. The key to the application of new media technology is the concept of display and content expression formed by the technology to meet the needs of museum education. The key to the application of new media technology is whether the display concepts and content expressions formed by the technology can meet the needs of museum education [13, 14].

In summary, the digital platform provides a platform for museum and visitor education. In short, digital platforms provide an important opportunity for a close connection between museums and visitors, especially as museums are becoming more popular with bring-your-own-device displays. In particular, museums are providing services at a time when bring-your-own-device is all the rage, making museums and visitors form a one-to-one connection and meeting visitors’ needs. This allows museums to connect with visitors on a one-to-one basis, to meet the diverse needs of visitors, and to better serve previously neglected populations. Making museums the museum’s knowledge is more relevant, so that information is not isolated and is interconnected [15]. The hands-on exploration of learning the museum has provided an enabling tool in the learning process. It also further enhances the museum’s own impact and truly the educational function of the museum is realized.

2. Applications of Virtual Reality

From this exhibit, we can see that there is a lot of room for museums to develop in the application of virtual reality technology. This is to be elaborated from several aspects: basic research, display platform, and cultural communication [16].

The reason why museums can become an important platform for contemporary cultural life exhibition, activity, and communication is not only because of its natural function of having activity venues, but more importantly, its function of collecting, preserving, repairing, and researching. Virtual reality technology in the construction of virtual space is not fabricated out of thin air, need a large number of graphics, video, and data support, and the museum’s basic research plays a vital role [17]. At present, all museums are in a large number of exhibits of digital collection, due to the level of different objects different collection of data if open to the public, how to better present are important aspects of the museum exhibition. The exhibition hall through the data optimization construction can be continuously updated virtual display space. At the same time can also grasp a good heritage exhibits which data is available for public display which is required to be kept confidential. This is a key research that is not available to other industries and units. In this research process carried out on the virtual reality technology on the number, quality, density, and specifici
cation of digital acquisition are proposed to study the requirements. At the same time, virtual reality in the continuous development will not have an impact on the value of the cultural relics themselves; this can also be explored in depth.
in this exploration and research. This also puts a demand on the direction of basic museum research [18].

The use of natural user interfaces is best known by the Cleveland Museum of Art. Cleveland Museum of Art Museum has a 12.2-meter-wide collection wall, which is the largest multitouch screen in the United States, allowing for multiple visitors to move and select freely and intuitively. The wall incorporates detailed graphic information on more than 4,000 works of art in the collection, and the display automatically switches themes every 40 seconds. Visitors approaching the display will click to enlarge the image according to their interests, and by looking at the large high-resolution image in detail, they will get more detailed information about the exhibit and its actual location in the museum, while the surrounding images will create a chain reaction that will show exhibits of similar themes, related artists, and similar time periods [19].

The natural user interface in our museums is mostly interactive for guidance and explanation such as introducing the background of the exhibition screens, and the Palace Museum is at the forefront of this application area. Many visitors are used to viewing calligraphy exhibits, and the “Digital Calligraphy Desk Copying the Lanting Preface” exhibit is based on this characteristic of the public, displaying the Lanting Preface on three high-definition screens and allowing visitors to randomly open the text in the exhibit on a surface tablet and write it with a “digital brush.” The characters written by themselves will be integrated into the original work and compared with it. This allows visitors to enjoy the heirloom masterpieces without barriers while also experiencing them with their hands, making the arcane calligraphy exhibits more relevant to modern visitors.

Experiential interactive displays are used in museums to create a variety of interactive effects. Visitors can establish an inquiry-based learning model through interaction, selecting the information they are interested in, actively thinking about it, discovering it, and feeling it. The interactive displays can be used to create a variety of interactive effects [19]. It can also be used to stimulate fantasy and curiosity and to deepen the understanding of the collection. It changes the way museums traditionally present their collections, so that instead of just staring at the exhibits and looking at the explanatory notes, visitors can use a series of simple actions to inspire them.

A series of simple manipulations that inspire visitors to explore the information in these exhibits, to discover which parts of the exhibition are interconnected, and to build a relationship with visitors directly. The exhibition is interconnected, building a direct interactive relationship with visitors, deepening their feelings of visiting, and making visitors take the initiative to obtain a new mode of interaction is created.

3. Kinect Somatosensory Technology Introduction

Kinect [20] is a body-sensitive input device for Xbox 360, a game console from Microsoft, with instant motion capture, image recognition, microphone input, speech recognition, and community interaction. The RGB camera is used to acquire 640 × 480 color images for skeletal tracking of human body keys. 3D depth sensors include infrared emitters and infrared COMS cameras. The microphone array is used for noise reduction and voice recognition. And Kinect V2 can acquire 25 skeletal points, 5 more than the previous generation: head, left fingertip, right fingertip, left thumb, and right thumb. Its color is 1080P, the depth camera is 512 × 424, and it can recognize the bones of 6 people with stable recognition and high accuracy.

How it works is that Kinect detects and captures the user’s gestures with the PrimeSense software and camera and then compares the captured images with its own internal human model. Each object that matches the existing human body model is created as a skeletal model, which is then converted into a virtual character that is triggered by recognizing key parts of the human skeletal model (as shown in Figure 1).

The Kinect program flow (shown in Figure 2) consists of five steps: initialization, image acquisition, analysis of tracking status, image display, and shutdown. The KinectGetDepthImage () and KinectGetSkeleton () functions obtain data frames from the color stream, depth stream, and skeleton stream and convert them into image types. Analyze the tracking status corresponding to Kinect-Judge Track () to determine the current tracking status and the control of the indicated situation. The display image corresponds to KinectDraw-Skeleton () and the display part of the main function, which displays the color map, depth map, and processing in real time displays the color map, depth map, and processed skeleton map [21].

With the rapid development of science and technology, the popularity of information technology and intelligence has brought disruptive changes to people’s lifestyles, and science and technology museums should pay more attention to user experience and exhibit interaction. Using the latest technology to improve and innovate the exhibits effectively will bring better user experience and will surely win more audience’s love. The author uses Unity3D as the engine to build an interactive platform and uses C# language as a bridge to connect Kinect, webcam, and other interactive devices, combined with Kinect augmented reality technology to innovate the model of science and technology exhibition hall exhibits and realize a prototype system of human-bird interaction [22].

The design of the Kinect-based human-bird interactive system is mainly completed by three functional modules interaction; the principle is shown in Figure 3.

Signal acquisition: using body sensors to record images and capture audience movements.

Signal processing: the collected data is analyzed and the resulting data is used to generate a virtual image corresponding to the real scene and to superimpose the fusion.

Image display: the final image is presented to the audience using a monitor with sound effects.

3D modeling is used to build the scene and the bird model. Before entering Unity, the character has to prepare basic actions, such as flying, winging, and standing by. There are three main types of character assembly and control provided by Unity: legacy, generic, and humanoid. Before importing the character into Unity, you need to design the FBX action file according to its different assembly modes in a targeted way.
After that, save the 3dsMax prepared bird character model to the Assets folder of your Unity project.

4. CL-KCF Algorithm

Since KCF does not deal well with target deformation scenes and occlusion scenes, does not track accurately when the target changes in size, and may cause subsequent tracking events during the long-term tracking process due to model error updates in a few frames, we propose our own CL(aose)-KCF algorithm based on the KCF algorithm. Therefore, based on the KCF algorithm, we propose our own CL(aose)-KCF algorithm, which is optimized in the following aspects.

1. The prediction of KCF is divided into two processes: location prediction and scale prediction, in which the location prediction is combined with a color histogram-based tracking model to effectively locate the target of deformation, and the scale prediction is added with a one-dimensional correlation filter to sample different scales around the predicted location to obtain the maximum scale response

2. The update strategy of KCF is improved by not only updating the model with the prediction results of the latest frame but also saving the results with high confidence from previous tracking into the sample sequence to reduce the effect of single-frame prediction bias.

3. Confidence verification of the KCF tracking results is performed, and when the confidence level is lower than the set threshold, the target position is re-detected by the redetection algorithm, and the tracking model is reupdated [23].

4.1. Location Forecast. In the KCF kernelized correlation filter, the target feature extraction is based on the HOG feature, which is able to extract the spatial structure of the target object under complex lighting changes and motion blur scenes, but the object spatial structure changes with the deformation of the object. Hog feature has poor tracking effect when dealing with such deformation scenes. However, the color distribution of the image does not change greatly in the deformation scenes and is relatively stable. The color feature handles well when the shape changes, but it is severely affected when the lighting changes discontinuously, and the color distribution alone is not sufficient to distinguish the target from the foreground [24]. In this section, a model synthesis of two complementary factors is used to handle complex scenes with color changes and object deformations. To be able to maintain the speed of real time, two independent both based on ridge regression tracking models are combined to achieve higher accuracy by combining the position detection scores of both models in a single training process with intensive sampling. For both models with the same regression function, their algorithm scores can be structurally similar in terms of dimensionality and predictiveness of results, and the higher scores can be used to identify more reliable target locations [25].

To deal with the deformation and color change of the object, we decided to use a color histogram model. For a single image, the color histogram tracking model needs to be trained with a series of square samples $x$ and the corresponding labels $y$. Similar to the correlation filtering, the histogram weight vector can be trained by minimizing the regression error function:

$$
\min_{m} \sum_{x} \left( \sum_{u \in R} m^T \varphi(x; u) - y \right)^2 + \lambda \|m\|^2, \quad (1)
$$

where $m$ is the histogram model weight vector, and the representation is the feature pixel values of sample $x$ within a defined region $R$. Although the minimization regression function of the histogram model is trained similarly to the error function of the correlation filter, the histogram weight vector cannot be cyclically shifted like the cyclic matrix in the correlation filter. The computational effort of $\varphi(\cdot)$ for a feature transformation with $M$ number of channels increases geometrically with the number of feature channels. Since the histogram feature score can be considered as an even-weighted election, a linear regression function can be applied to each

![Figure 1: Kinect working principle.](image1)

![Figure 2: Kinect program flow.](image2)
By applying the solution of ridge regression, the above minimum loss function is solved to obtain:

$$m^i = \frac{p^i(0)}{p^i(0) + p^i(S) + \lambda}.$$  

For each dimension in, $p^i(R)$ is the $j$th element of the vector $p$, which represents the proportion of nonzero features $j$ in the pixels in the region $R$. After obtaining the histogram weight vector, in a given image sample $z$, we are able to compute the score key-value pairs for each pixel color in $m$:  

$$f_c = m^T \varphi_{set(z)}.$$  

And the dense histogram response map is obtained by integrating the image, for the online update of the color histogram model for:

$$P_j(0) = (1 - \eta_j)P_{j-1}(0) + \eta_j P'_j(0),$$

$$P_j(S) = (1 - \eta_j)P_{j-1}(S) + \eta_j P'_j(S).$$

$\eta_j$ refers to the adaptive update rate of the color histogram model, $P_j(R)$ is a vector of $P'_j(R)$, $j = 1, 2 \cdots m$.

4.2. Model Association. After combining the spatial structure features of the object and the color histogram features, it can reduce the reliance on the spatial structure features of the object and also reduce the situation that it is difficult to distinguish the object color when it is close to the color of the surrounding environment by simply taking the color features. The kernelized correlation filter tracker by the template-like approach is then able to distinguish target objects in complex environments well, while the color histogram model based on the color probability distribution is able to improve the probability that the kernelized filter loses the object target in the case of sharp deformation of the object, thus improving the robustness of both tracking models. This section completes the joint algorithm by calculating the tracking scores of the two algorithms:

$$S = \mu_h S_h + \mu_s S_c,$$

where $S_c$ refers to the correlation response value score of the kernelized correlation filter in the image, $S_h$ refers to the response value score of the color feature histogram-based tracking model in the image, and $\mu_h + \mu_s = 1$.

$$S_h = \omega^T \left( \frac{1}{|C|} \sum_{r \in C} I[r] \right).$$

In the score formula of the color histogram, $C$ refers to the candidate sample region in the image, while $r$ refers to the pixel values within the region, $\omega$ refers to the weight vector, and $I$ denotes the target feature based on the color histogram. Finally, the score of the color histogram is obtained by calculating the average score.

$$S_c = f(z) = F^{-1} \left( k^T \Theta \hat{\alpha} \right).$$

In the score formula of kernelized correlation filtering, it is the score obtained by cyclic shifting of the candidate sample set obtained from the base samples, by performing the correlation response of the image for each candidate sample. Where $F^{-1}$ refers to the inverse transform of the discrete Fourier and clothing refers to the spatial domain solution of the optimization parameters [12].

For solving the weights called $h$ and $\eta$ of the two tracking algorithms, the dynamic weighting approach is used here by comparing the highest values of the response values of the two tracking algorithms and then solving:

$$H_h = \frac{\max (S_h)}{\max (S_h) + \max (S_c)}.$$  

4.3. Scale Prediction. The process of fusing scale changes can be divided into two parts, location detection and scale detection, depending on the purpose of tracking. Position detection
is also the workflow of KCF kernelized correlation filtering that we described in Section 4.1. By densely sampling near the target object position by cyclic shifting, after extracting HOG features, the kernelized correlation filter of the object is obtained by converting the discrete Fourier transform from a low-dimensional convolution calculation to a high-dimensional dot product operation, training the optimal error function for the kernelized correlation filter to learn the model is determined by the response map of the filter in a new frame of the image with the maximum response at the 0 marker position. In contrast, the scale detection is based on the location of the target determined by the location prediction, and multiscale sample sampling is performed around this location to train the scale filter to find the largest correlation response in the scale as the final scale of the object, and then the model parameters of the filter are updated. The method framework of scale prediction is shown in Figure 4.

We additionally introduce a one-dimensional scale filter to detect and update the scale changes while performing the location prediction kernelization-related filter updates. For the candidate samples obtained by sampling different sizes around the predicted target location, we unify the samples of different sizes into the size of the target initial frame by using a bilinear interpolation method, feature extract these candidate samples on the HOG gradient histogram, and train the optimal error function to obtain our scale update filter. In the image frame, the size of the base sample of the scale filter is $s \times l$, the size of the object frame in the initial frame is $M \times N$, and we sample for samples of different sizes around the target object as the size frame of $m \times n$, $m = r^n M$, $n = r^n N$, with the scale parameter $r$.

The equation for the response value $\text{Res}$ of the target sample at different scales is

$$\text{Res} = \mathcal{F}^{-1} \left\{ \frac{\sum_{k=1}^{d} j_{F} \bar{M}^k}{N + \lambda} \right\},$$

where $d$ is the feature dimension; $\lambda$ is the regularization variable to prevent overfitting; $\bar{M}^k$ is the complex conjugate of $M^k$, and the size of the scale corresponding to the largest response value is used as the target scale of the current image.

5. Experimental Results and Analysis

5.1. Phantom Imaging Technology. Now there are also many museums and memorials using phantom imaging technology for exhibition displays. For example, the West Han Yangling Museum uses phantom imaging to switch multiple physical scenes and match the image and sound effects to present scene of the civil and military officials going to court during the reign of Emperor Jing of Han Dynasty (Figure 5) is very graphic.

The China Mintai Yuan Museum also incorporates phantom imaging technology in its displays to showcase the common customs of Fujian and Taiwan in terms of life rituals and food culture and to complement the surrounding physical displays to create a strong regional atmosphere, making it easier for visitors to appreciate the closeness of the two peoples and to empathize emotionally.

In addition to short videos and videos of the collection, the painting and calligraphy collection is also well suited for display using giant screen projection technology. The 100-meter long scroll was created based on Zhang Zeduan’s version of “Qingming Shanghe Tu” in the Song Dynasty and was stitched together and fused with 12 cinema-grade projectors, magnifying the original work about 30 times and showing it on a giant projection screen of 6.5 meters in height and 128 meters in length, recreating scenes from the life of citizens in the Northern Song Dynasty. Under the rendering of surround sound, the dynamic people, and scenery in the projection, visitors seem to be walking in the streets of Bianjing in an immersive way [14].

In addition, the giant dynamic version of “Qianlong’s Southern Tour” of the National Museum of China in 2014, using 3D modeling technology, running track setting, and edge fusion technology, presented the first volume of “Qi Bijing Shi” dynamically on a giant screen 30 meters long and 4 meters high, vividly restoring the magnificent historical situation in the painting and recreating the grandeur and splendor of the Kang and Qian dynasties. Visitors can see not only the movements of the Qianlong emperor through the digital display but also the dress of the civil and military officials, the rituals of the tour, the customs of the city, and even the natural geography and humanistic landscape between the north and south of the river, thus gaining a more comprehensive and in-depth experience of the charm of this painting!

5.2. Location Forecast. In order to verify that the combined use of the kernelized correlation filter tracking algorithm and the color histogram-based tracking algorithm can better handle scenes in which the target undergoes drastic deformation, we selected four video sequences from the standard video sequences of OTB-50 in which the target deformation is the main factor, and supplemented with complex background, fast movement, out-of-field, and illumination changes, etc., to qualitatively The effectiveness of this strategy is illustrated from both qualitative and quantitative perspectives. The qualitative perspective is based on the tracking effect of the improved KCF algorithm and the combined CL-KCF algorithm, while the quantitative perspective is based on the pixel error at the
The list of joint test video sequences is shown in Table 1.

In the video sequence of bird1, the bird keeps changing its form as it flies forward with its wings. In frames 11 and 15, both tracking algorithms still maintain good tracking results, but in frame 29, as the bird’s wings beat downward, the tracking algorithms both show a certain degree of drift. After frame 30, with the accumulation of KCF tracking algorithm errors, the original bird-based tracking model drifts and can only maintain the tracking of the wings subsequently, while our joint algorithm is still able to maintain the target tracking frame in the bird’s body to maintain good tracking results.

From the center pixel error plot in Figure 6, we can see that the average pixel error of our CL-KCF algorithm compared to KCF decreases from 152.34 to 102.50, a decrease of nearly 33%, and the success rate of tracking in the success rate plot is within the center pixel error of 100 or less, a significant improvement.

In Figure 7, we can see that our CL-KCF algorithm fluctuates within a pixel value of 10 in the central pixel plot, while the KCF algorithm rises quickly and linearly above 100, confirming that the KCF algorithm has already lost the athlete at the beginning of the run. Similarly, in the success rate graph, we can see that our algorithm has achieved a 100% success rate of tracking at a pixel threshold of 20, which illustrates the advantage of our improved algorithm.

From the successful coverage graph in Figure 8, it is obvious that our CL-KCF algorithm covers the car much better than the KCF algorithm in the interval from frame 100 to frame 600 when the car size changes. The average coverage of CL-KCF is 0.75 while KCF is only 0.48, which

| Video sequence | Number of frames | Complex scene factors                     |
|----------------|------------------|------------------------------------------|
| Bird 1         | 408              | Object deformation, rapid movement, field of view |
| Bolt 2         | 293              | Object deformation, complex background    |

**Figure 5: Phantom imaging at the Hanyangling Museum.**

**Table 1: Model joint test video sequence list.**

**Figure 6: Bird1 video sequence of central pixel error plot and success rate plot.**

(a) Center pixel error map

(b) Success rate chart
is a 27% improvement in coverage, due to the adjustment of the target prediction frame size, which is closer to the actual size of the target in the frame.

6. Conclusions

Virtual reality technology is a direction in digital technology, compared to other digital means, virtual reality technology has a strong ability to restore time and space, with other exhibits do not have the immersive experience feeling. Combining Kinect augmented reality technology with science education and effectively improving and innovating exhibits will certainly enhance the interactivity and fun of science exhibits and help popularize scientific knowledge widely and deeply, thus realizing the change from "passive science popularization" to "active science popularization" transformation. These advantages make its function in the basic research, and display platform and cultural dissemination of museums produce great effect.

Data Availability

The experimental data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declared that they have no conflicts of interest regarding this work.

References

[1] L. Reitstätter, H. Brinkmann, T. Santini et al., “The display makes a difference: a mobile eye tracking study on the perception of art before and after a museum’s rearrangement,” Journal of Eye Movement Research, vol. 13, no. 2, 2020.

[2] P. Fontoura and M. Menu, “Visual perception of natural colours in paintings: an eye-tracking study of Grünewald’s Resurrection,” Color Research & Application, vol. 46, no. 3, pp. 582–594, 2021.

[3] D. Christidou and L. Reitstätter, “Trails of walking-ways of talking: the museum experience through social meaning mapping,” in In international conference on emerging technologies and the digital transformation of museums and heritage sites, Springer, Cham, June 2021.

[4] S. Blanco-Pons, B. Carrión-Ruiz, M. Duong, J. Chartrand, S. Fai, and J. L. Lerma, “Augmented reality markerless multi-image outdoor tracking system for the historical buildings on Parliament Hill,” Sustainability, vol. 11, no. 16, p. 4268, 2019.

[5] S. Alletto, R. Cucchiara, G. Del Fiore et al., “An indoor location-aware system for an IoT-based smart museum,” IEEE Internet of Things Journal, vol. 3, no. 2, pp. 244–253, 2016.

[6] A. P. Plageras, K. E. Psannis, C. Stergiou, H. Wang, and B. B. Gupta, “Efficient IoT-based sensor BIG Data collection-
processing and analysis in smart buildings,” *Future Generation Computer Systems*, vol. 82, pp. 349–357, 2018.

[7] M. Mohammadi, A. Al-Fuqaha, M. Guizani, and J. S. Oh, “Semisupervised deep reinforcement learning in support of IoT and smart city services,” *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 624–635, 2018.

[8] M. Jia, A. Komelj, Y. Wang, and R. Srinivasan, “Adopting internet of things for the development of smart buildings: a review of enabling technologies and applications,” *Automation in Construction*, vol. 101, pp. 111–126, 2019.

[9] J. Siryani, B. Tanju, and T. J. Eveleigh, “A machine learning decision-support system improves the internet of things’ smart meter operations,” *IEEE Internet of Things Journal*, vol. 4, no. 4, pp. 1056–1066, 2017.

[10] X. Guo, N. Ansari, L. Li, and H. Li, “Indoor localization by fusing a group of fingerprints based on random forests,” *IEEE Internet of Things Journal*, vol. 5, no. 6, pp. 4686–4698, 2018.

[11] C. Zhang, T. Xie, K. Yang et al., “Positioning optimisation based on particle quality prediction in wireless sensor networks,” *IET Networks*, vol. 8, no. 2, pp. 107–113, 2019.

[12] S. Nižetić, P. Solić, D. L. D. I. González-de, and L. Patrono, “Internet of things (IoT): opportunities, issues and challenges towards a smart and sustainable future,” *Journal of Cleaner Production*, vol. 274, p. 122877, 2020.

[13] P. Spachos and K. N. Plataniotis, “BLE beacons for indoor positioning at an interactive IoT-based smart museum,” *IEEE Systems Journal*, vol. 14, no. 3, pp. 3483–3493, 2020.

[14] S. Ma, Q. Liu, and P. C. Y. Sheu, “Foglight: visible light-enabled indoor localization system for low-power IoT devices,” *IEEE Internet of Things Journal*, vol. 5, no. 1, pp. 175–185, 2018.

[15] H. Che, J. Wang, and A. Cichocki, “Bicriteria sparse nonnegative matrix factorization via two-timescale duplex neurodynamic optimization,” *IEEE transactions on neural networks*, vol. PP, pp. 1–11, 2021.

[16] H. Che and J. Wang, “A two-timescale duplex neurodynamic approach to mixed-integer optimization,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 1, pp. 36–48, 2021.

[17] H. Che and J. Wang, “A two-timescale duplex neurodynamic approach to biconvex optimization,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 8, pp. 2503–2514, 2019.

[18] H. Graf, J. Keil, A. Pagano, and S. Pescarin, “A contextualized educational museum experience connecting objects, places and themes through mobile virtual museums,” *Digital Heritage*, vol. 1, pp. 337–340, 2015.

[19] N. Ghouaiel, S. Garbaya, J. M. Cieutat, and J. P. Jessel, “Mobile augmented reality in museums: towards enhancing visitor’s learning experience,” *International Journal of Virtual Reality*, vol. 17, no. 1, pp. 21–31, 2019.

[20] S. Karaman, A. D. Bagdanov, L. Landucci et al., “Personalized multimedia content delivery on an interactive table by passive observation of museum visitors,” *Multimedia Tools and Applications*, vol. 75, no. 7, pp. 3787–3811, 2016.

[21] C. Matuk, “The learning affordances of augmented reality for museum exhibits on human health,” *Museums & Social Issues*, vol. 11, no. 1, pp. 73–87, 2016.

[22] J. M. Harley, E. G. Poitras, A. Jarrell, M. C. Duffy, and S. P. Lajoie, “Comparing virtual and location-based augmented reality mobile learning: emotions and learning outcomes,” *Educational Technology Research and Development*, vol. 64, no. 3, pp. 359–388, 2016.

[23] M. Rainoldi, B. Neuhofer, and M. Jooss, “Mobile eyetracking of museum learning experiences,” *Information and communication technologies in tourism*, vol. 2018, pp. 485 pages, 2018.

[24] E. G. de Ravé, F. J. Jiménez-Hornero, A. B. Ariza-Villaverde, and J. Taguas-Ruiz, “DiedricAR: a mobile augmented reality system designed for the ubiquitous descriptive geometry learning,” *Multimedia Tools and Applications*, vol. 75, no. 16, pp. 9641–9663, 2016.

[25] J. Amakawa and J. Westin, “New Philadelphia: using augmented reality to interpret slavery and reconstruction era historical sites,” *International Journal of Heritage Studies*, vol. 24, no. 3, pp. 315–331, 2018.