Searching for places or objects in a collection of photographs that are spatially related is a task that is becoming more and more relevant in today’s digital world. Photographers, designers, and hobbyists must often search through thousands of photos to find what they are looking for. People search for places visually through applications such as Google Street View or Microsoft PhotoSynth. Furthermore, collaborators are increasingly using images for integrating the spatial world into remote collaboration. In this position paper, we approach spatial search from the viewpoint of virtually navigating a remote world represented by a set of spatially related images for the ultimate purpose of improved telecollaboration. Telecollaboration is an enhanced form of videoconferencing that allows participants to collaborate in more ways than just seeing and hearing each other. We first describe related work in image-based spatial search, then present our recent research in applying image-based spatial search to telecollaboration, and finally give our thoughts on some current challenges and directions for visual spatial search.

Searching for images associated with physical locations has been approached in various ways in the past. Using solely automatic GPS annotations and manually added keywords has long been the naïve approach to searching through such images that are tied to a spatial location. More recently, 2D image features (such as SIFT features) and visual vocabularies have been used to create systems that match locations or objects across videos or photographs [1]. In addition, research in structure from motion and stereo vision (and also vision with depth sensors) has allowed computers to analyze spatially related photos to recover the 3D relation between them and the 3D structure of the commonly observed scene [2]. Researchers have also been analyzing image-based localization—that is, determining the geo-location of a single photograph; this is useful for mobile localization when GPS is often unavailable [3].

Recently we in the Four Eyes Lab at UCSB have been researching ways to integrate the physical environment into telecollaboration [4]. In our case, a local user holds a tablet device which transmits a video stream to a remote user. The remote user is then able to spatially annotate the local user’s scene by using augmented reality. Furthermore, the remote user can virtually navigate or search through the scene by using a subset of the video stream frames as “key-frames” that encompass different viewpoints of the scene’s spatial area. By using augmented reality and image-based navigation, our prototype telecollaboration application allows users to virtually annotate the remote physical world and to search through it in a visual
way. Our results have shown that users prefer our user interface over more traditional telecollaboration interfaces.

In this context, we have been exploring the challenges of virtual navigation of a remote scene for telecollaboration. Virtual navigation typically encompasses three tasks—exploration, search, and inspection. To investigate how different user interfaces affect user performance, we used a naïve search task in an exploratory user study and found insights into designing future user interfaces for image-based navigation of virtual environments [5]. Later, we expanded our interface to use a touchscreen and found that users prefer our interface over others [6].

By investigating virtual navigation for telecollaboration, we have unearthed aspects and challenges of visual spatial search that have not been explored much in the past. This includes the real-time collaborative aspect of visual spatial search. Previous systems, such as Google Street View or Microsoft PhotoSynth, must build offline a virtual scene model for users to spatially search through; our research, however, brings in a real-time element to building this model through which users can spatially search for items of interest.

Another challenge for visual spatial search in the context of telecollaboration is the cognitive aspects of designing the user interface. In constrained situations, such as Google Street View, clicking in predefined areas (e.g., roads) makes it easy to search and navigate through images with relatively low cognitive load on the user. In unconstrained situations, however, current visualizations for image-based navigation interfaces are not sufficient and cause high cognitive load on the user [5]. On the one hand, users can simply scroll through lists of photos in a brute force manner with low cognitive load but with also extremely low throughput. On the other hand, the system may be able to index visual features common to multiple images [1] or utilize 3D structure [4] to increase search throughput while maintaining a low cognitive load.

Perhaps the two main challenges currently for visual spatial search are what some researchers call “perceptual aliasing”—when very similar visual information appears in multiple places, making it hard to distinguish between them—and “image variability”—when the visual appearance of a certain place changes temporally, making it hard to recognize the same place over time. For example, it is hard to geo-localize a photo of a bicycle without any contextual information and it may be hard to recognize the same park from photos taken during different seasons, such as winter and summer. However, for small areas, such as might occur in telecollaboration, identifying the relative associations between places may be more achievable. All of these considerations lead us to interesting research questions, such as, “What is the relationship between searching through a collection of images and navigating/exploring a remote spatial environment?” And, “Is it possible to juxtapose research advances in image-based search and 3D scene reconstruction, along with human computer interaction techniques, to enable efficient and intuitive visual spatial search?”

In conclusion, we have begun to explore visual spatial search in the context of telecollaboration, and this has opened up important questions in how to apply previous research advances in visual search to the current challenges of realizing visual spatial search. As research continues to progress towards addressing the spatial component of visual search, we will begin
to see systems that enable humans to more intuitively and efficiently search through both the visual and spatial realm.

Works Cited


