

Technical Disclosure Commons

Defensive Publications Series

January 2021

Image Region-Snapping Based on Touch-and-Hold and Drag Gestures

Brett Aladdin Barros

Paulo Coelho

Follow this and additional works at: https://www.tdcommons.org/dpubs_series

Recommended Citation

Barros, Brett Aladdin and Coelho, Paulo, "Image Region-Snapping Based on Touch-and-Hold and Drag Gestures", Technical Disclosure Commons, (January 04, 2021)

https://www.tdcommons.org/dpubs_series/3939



This work is licensed under a [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/).

This Article is brought to you for free and open access by Technical Disclosure Commons. It has been accepted for inclusion in Defensive Publications Series by an authorized administrator of Technical Disclosure Commons.

Image Region-Snapping Based on Touch-and-Hold and Drag Gestures

ABSTRACT

Current techniques for image-based search perform a search the entire image and do not offer an easy to use mechanism for seamless selection of portions of an image. This disclosure describes techniques to automatically suggest refined (subset) regions within a given image that can be used to trigger a search. In response to a gesture, e.g., touch-and-hold, a sub-region within the image that includes an object is identified and is shown within a bounding box. In response to a discontinuation of the gesture, e.g., lifting of the user's finger, a semantic search is performed for content that is similar to the contents of the bounding box. Other types of operations can also be performed specific to the bounding box.

KEYWORDS

- Image recognition
- Image segmentation
- Semantic image analysis
- Touch and hold
- Scene understanding
- Visual search
- Bounding box
- Region-snapping

BACKGROUND

On a mobile phone or other touchscreen device, a user can touch and hold (T&H) on text to select it, then search for it. Similarly, a user can T&H an image to select and search for it. However, in certain situations, a user requirement is to search a particular part of an image, e.g., the suit a person depicted in the image is wearing, not including the person or the scene in which the person is situated. Current techniques for image-based search perform a search for the entire image and do not offer an easy to use mechanism for seamless selection of portions of an image.

DESCRIPTION

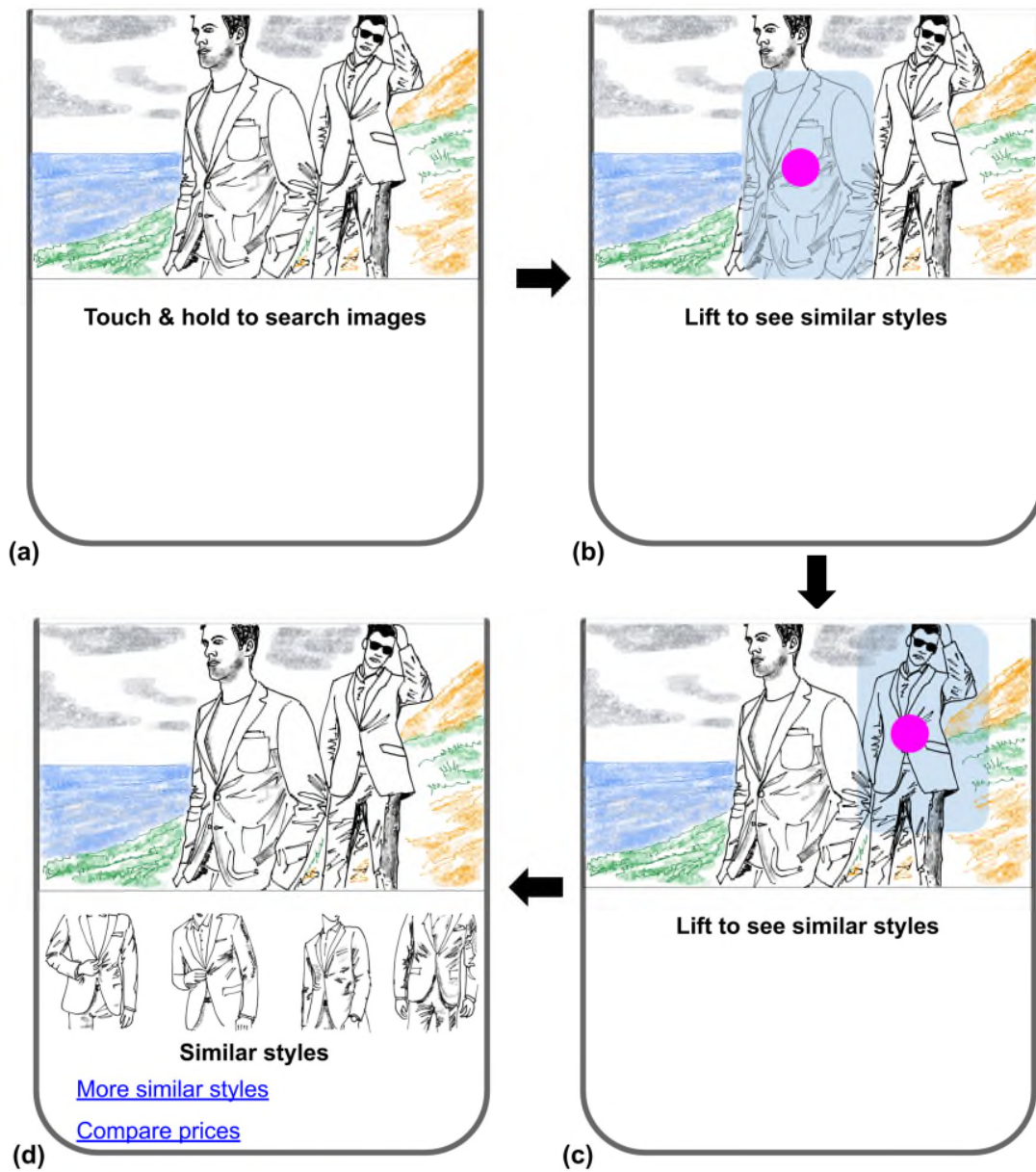


Fig. 1: Using touch, hold, and drag to snap to a region of an image

Fig. 1 illustrates the use of a touch-and-drag gesture to snap on to a region of an image. In Fig. 1(a), a user views an image on a touchscreen device. The user performs a touch-and-hold gesture, in response to which a section of the image is grabbed, as shown in Fig. 1(b). Note that the whole image is not grabbed; rather, only a relevant region close to the point of

touch-and-hold is grabbed. The criteria for region selection are elaborated in greater detail below. The user then drags the finger across the screen, which triggers a different region of the image to be selected, as shown in Fig. 1(c). When the user lifts their finger, an automatic search for images similar to the selected section of the image is performed, as shown in Fig. 1(d).

The automatic selection of a subset region of the larger image can be performed as follows. When a user performs a touch-and-hold (T&H) gesture at a position on the image, the object (container) that encapsulates the touch coordinates is detected. Some criteria for identifying the encapsulating visual object include: centricity (how close the touch is to the center of the container), color blocks, high contrast edges, etc. Suitably trained machine learning models can be utilized to detect object boundaries.

When a visual object is detected, a bounding box preview, as shown in Fig. 1(b) and Fig. 1(c), is displayed to inform the user of the area auto-selected in response to the touch-and-hold gesture. The auto-selected area can be modified by the user, e.g., by dragging their finger to a different portion of the image. If the user's finger position goes beyond the container and nearer to the center point of a second detected object, the second object is selected. The lifting of the user's finger (after the touch-and-hold or drag), is an indication of the intent to search. The object within the bounding box is semantically analyzed. The resultant semantic understanding is utilized to trigger a search for content that is similar to the contents of the bounding box. While Fig. 1 illustrates a search, any suitable operations can be performed based on the automatically selected sub-region of the image.

In order to distinguish between a regular touch and hold gesture as implemented in the application from the gesture described herein, one or more of the following additional criteria can be applied:

- Determine that the swipe is from a region of the screen, e.g., originating at an edge, or on a system navigation button, followed by leaving the finger on the screen to establish a touch and hold.
- Determine that the touch and hold is performed within a particular system or application state, e.g., only when a virtual assistant has been invoked/ is active

CONCLUSION

This disclosure describes techniques to automatically suggest refined (subset) regions within a given image that can be used to trigger a search. In response to a gesture, e.g., touch-and-hold, a sub-region within the image that includes an object is identified and is shown within a bounding box. In response to a discontinuation of the gesture, e.g., lifting of the user's finger, a semantic search is performed for content that is similar to the contents of the bounding box.