

# Directed Manipulation With Respect To Focal Rings

Darren Edge<sup>1</sup>

<sup>1</sup>Microsoft Research Asia  
5F Beijing Sigma Center, 49 Zhichun Road  
Beijing, China  
Darren.Edge@microsoft.com

Xianghua Ding<sup>2,1</sup>

<sup>2</sup>Department of Informatics  
University of California, Irvine  
Irvine, USA  
Xianghua.Ding@gmail.com

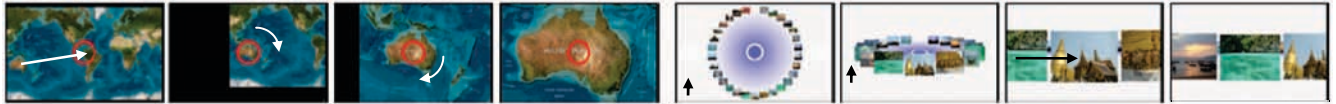


Figure 1. Interaction with focal rings. Left – a focal ring provides a Polar reference frame for interacting with a map: drag target in the direction of the focal ring to pan; circle around ring to zoom. Right – a focal ring provides a Cartesian reference frame for interacting with a carousel: stroke vertically to tilt the carousel; stroke horizontally to rotate the carousel about the focal ring.

## ABSTRACT

In this paper we describe the notion of a *focal ring* interface: an interface comprising a centrally-located graphical ring that provides a visual focus for interaction with respect to itself and its associated representation (such as a map, media collection, or information hierarchy). Our focal rings provide an opportunity to assign independent effects to the two dimensions of a display surface, creating *implicit gestural modes* (e.g. navigation versus scaling) that can be dynamically switched between in the context of a single continuous stroke. A focal ring can define the origin of a Polar or Cartesian frame of reference, interpreting touch gestures directed either through and around, or horizontal and vertical to itself. We illustrate the potential for such *directed manipulation* with respect to focal rings through the design and evaluation of ring-centric interfaces based on a variety of physical metaphors.

## Author Keywords

Directed Manipulation, Focal Rings, Touch Interaction, Implicit Gestural Modes

## ACM Classification Keywords

H5.2 User Interfaces: Interaction styles

## INTRODUCTION

This paper presents an alternative to interaction by direct manipulation [15] or by gesturing: *directed manipulation* with respect to focal controls. By introducing a ring-shaped control to the centre of an interaction space – defined by the physical extent of a smaller display device (e.g. the screen of a mobile phone or tablet PC) or by a virtual partition of a larger surface (e.g. an interactive

tabletop or wall) – we can suggest a new frame of reference for touch interactions. Figure 1 gives examples of such *focal ring* reference frames for interaction with a map and a carousel respectively.

With respect to a focal ring, finger strokes can be interpreted differently when directed towards or away from, into or out of, through or around, or horizontal or vertical to the ring. This gives rise to *implicit gestural modes*: modes that can be dynamically switched between in the context of a single continuous stroke, by varying the direction of the stroke relative to the focal ring. This approach to interaction combines elements of both direct manipulation and gesturing, and was found to support a style of interaction that is easy to learn, use, and explain.

In the remainder of this paper, we describe the origin of the focal ring concept, our exploration of the design space, and an analysis of focal rings from a variety of perspectives. We then describe our implementations of three ring-centric interfaces, for interaction with large images, media collections, and information hierarchies respectively. We conclude with a presentation of results from a user study.

## ORIGIN

The inspiration for our focal ring concept arose from our collaboration with the Chinese television company CCTV, concerning the potential use of a touch interface on their Olympic Channel's nightly program during the Beijing 2008 Olympic Games. We saw this as a unique opportunity to appropriate the Olympic symbol<sup>1</sup> as an interactive control, in particular by developing an interface based on the concept of focal Olympic rings.

© The Author 2009.

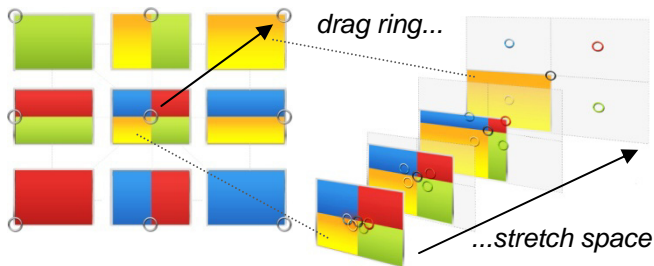
Published by the British Computer Society

<sup>1</sup> The Olympic symbol of five interlaced rings and the Olympic identifications (e.g. Olympics, Olympic Games) are the exclusive properties of the International Olympic Committee. Limited usage rights were granted to CCTV in their role as official broadcaster of the Beijing 2008 Games.



**Figure 2. Stretchable quad UI on CCTV Olympic Channel.**

The first part of our design proposal was to locate the Olympic symbol of five interlocked rings in the centre of the interface, and to divide physical display space into four quadrants by extending crosshairs from its central, black ring (Figure 2). This ring acts as a draggable control that stretches the virtual “quad”, such that individual or adjacent quadrants can be maximized to full screen (Figure 3). The potential benefit of this framework for TV use is that four concepts can be presented separately, in ways that are visually striking yet simple enough to be comprehensible when viewed on home audience TV sets.



**Figure 3. Design of stretchable quad control. Left: The 9 display configurations showing 1, 2, or 4 quadrants. Right: stretching the lower left quadrant to fill the whole display.**

Figure 3 also shows how the remaining four Olympic rings were integrated into this *stretchable quad* control. From an initial connected configuration at the centre of the display, as the black Olympic ring is moved towards the display’s boundary, the other four, coloured Olympic rings disconnect and move towards the centres of their respective quadrants. Once a ring has “homed” in on this central position (such as the yellow ring in Figure 3, right), it becomes a *focal ring* used to guide subsequent interactions at this level. The five rings thus engage in an iconic “dance”, evocative of the wings of a butterfly, whenever the *stretchable quad* is resized. On release of the black ring, it can either “snap” to the closest stable position (centre, corner, or mid-side), or “slide” to the next stable position in its direction of motion. Whilst the former was found to give a greater feeling of control, the latter minimized physical exertion when interacting with larger displays and was used in the final CCTV interface broadcast on-air (Figure 2).

Our *stretchable quad* control can be seen as a new and general way of presenting, comparing, and navigating between four different panels of content. However, it is still what we would call a direct manipulation interface – the black Olympic ring “control” is dragged by the user according to their desired multiplexing of the four content panels onto the single display space.

### From Five-Ring Framework to Four Focal Rings

This *stretchable quad* was the high-level presentation framework in our proposed design, capable of embedding four quadrants of lower level content. Since in the early stages of our collaboration with CCTV the exact nature of the TV program was yet to be decided upon, we illustrated the potential for faceted browsing (e.g. [5]) of Olympic media using a *stretchable quad* framework containing four *focal ring* quadrants. Three of these quadrants would be used to filter Olympic media by *time* (the schedule of 55 morning, afternoon, and evening sessions), *place* (the map of 40 venues), and *sport* (the 38 disciplines covering 302 events). The final *media* quadrant would then display the filtered set of media for the specified events (line-ups, results, photos, videos, etc.). We therefore designed four focal ring interfaces to interact with the map, schedule, structure, and media of Olympic events.

#### *Focal Ring “Telescope”*

In our first design, for panning and zooming a map of Beijing (or other large image such as a photo), a central ring floating above the map provides a focus for interaction that mirrors the importance of “centre” in such activities. It acts as both the origin of zoom operations and the destination for pan operations that precede and follow such zooming. By adopting a Polar frame of reference, such that strokes are only interpreted as pan operations when they are “targeted” towards, away from, or through the focal ring, “twisting” circular strokes around the ring are free to be interpreted as bidirectional zoom operations, analogous to the operation of a screw-threaded telescope (Figure 4a).

This *telescope* design assigns independent functions to the two dimensions of a Polar frame of reference. This can be seen as related to the control mapping of the OrthoZoom Scroller [3], which uses two Cartesian dimensions to independently manipulate the location and scale parameters of a one-dimensional information space. In contrast to such *dimension specializations*, the ChiralMotion technique [4] uses *dimensionality reduction* to collapse two-dimensional motion on a trackpad into a single dimension of navigational control. Finally, gestural techniques based on “rubbing and tapping” [12] as well as speed-dependent coupling between pan and zoom [8] support similar single-touch panning and zooming within a 2D visual representation such as a map, without relying on visible controls or explicit mode switching.

#### *Focal Ring “Tunnel”*

Our second design was another application of the Polar reference frame – a cylindrical “time tunnel” that vanishes at a distant focal ring, used for navigating the schedule of event sessions (Figure 4b). Its inner surface denotes both the “passage of time”, and intervals of it (sessions, days, chunks of days), with concentric circles. The user moves the tunnel relative to their viewpoint by stroking away from or towards the ring, “sliding” forwards or backwards in time respectively. A time interval is “stretched” by touching within it and circling around the ring.

This *tunnel* design uses a familiar perspective-based visualisation applied to the presentation of time. An earlier “time tunnel” and other perspective tunnel visualisations [9] used tunnels with square cross-sections, in contrast to our cylindrical, ring-centric design.

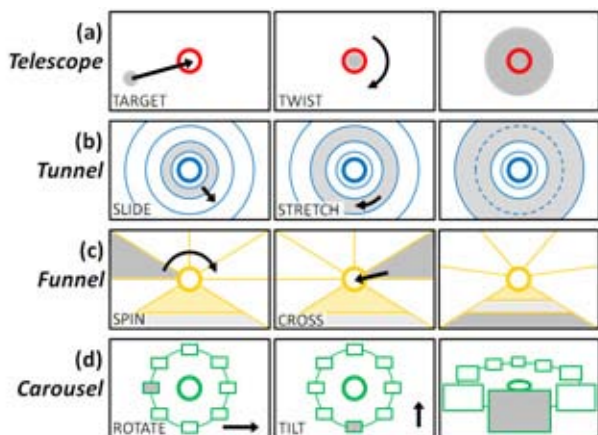


Figure 4. Initial focal ring designs.

#### Focal Ring “Funnel”

Our third design supports the rapid selection of any of the 302 Olympic events (e.g. “Men’s Individual Pursuit”), which requires navigation through the hierarchical tree structure of multiple sport types (e.g. “riding”) and disciplines (e.g. “track cycling”). Our solution was to represent “child” nodes as multiple sectors fanned out across the top three edges of the display, and the history of “parent” nodes as a stack growing upwards from the bottom edge of the display. The focal ring located at the centre of the display acts as a kind of “funnel” linking the two: dragging child nodes into the ring “pushes” them downwards, appearing at the base of the parent stack and triggering the display of the next generation of children (Figure 4c). Dragging out from the ring “pulls up” the children of the last node to be pushed into it.

This *funnel* design is a “goal crossing” interface (e.g. [1,2]) in which selection is through lines being drawn across targets, rather than targets being “clicked” or “tapped” in the conventional way. Our *funnel* most closely resembles an interface-wide FlowMenu [7], which makes similar use of a central “rest area” with crossing actions triggering transitions within a hierarchy. However, our *funnel* design is essentially the inverse of the FlowMenu and other marking menus: in our interface it is the last sector selected *on entry* to the focal ring that determines the effect, whereas the effect of crossing in traditional marking menus is determined *on exit* from the rest area.

#### Focal Ring “Carousel”

For our fourth focal ring interface, the need to support a variable-sized “result set” of media (filtered by prior selections of time, place and sport) made carousel browsing a suitable design choice. To accommodate both a focal ring and the unobscured full-screen display of media, we created a special “tilting” carousel (Figure 4d) that couples viewing angle with viewing distance. Upwards strokes anywhere on the display tilt the top half of the carousel away from the viewer. At the same time,

their viewing position moves closer, until the item at the base of the carousel fills the whole screen. Stroking left or right anywhere on the display rotates the carousel in the appropriate direction, while stroking down tilts the carousel back to a bird’s-eye view of all items. Once an item is “maximized”, items to either side can be stroked to directly in slideshow style.

This *carousel* design draws on previous work on carousel browsing [19], but additionally supports a continuous transition from the context of many items (bird’s-eye view) to the detail of a single item (full-screen view). Many approaches to managing detail+context in a limited space (e.g. fish-eye views) have been developed previously [17].

#### Discussion

Our original goal was to create new and engaging ways to interact with large touch-screen interfaces that would be broadcast to home television audiences during the Olympic Games. We succeeded in this goal, but our success was tempered in two respects. While the *stretchable quad* made it onto the show and was agreed by presenters and producers alike to support a visually distinctive presentation style, the program segment containing our interface was cut after the fourth night<sup>2</sup>. In practice, the way in which it was used to sequentially present four panels of largely similar content did not make full use of its fluid transitions and side-by-side comparisons. Similarly, while our four single-ring designs (Figure 4) were praised by the production team as easy to use and compelling to watch, they were not implemented for inclusion in the final interface. Ultimately, the pressure on the presenter to deliver scripted content in a fixed time period outweighed the desire to engage audiences with interactive presentations. However, our creation of these Olympic-inspired designs had convinced us of the value of focal rings as an interaction device, and motivated further exploration of the design space.

#### EXPLORATION

The focal ring designs of Figure 4 differ in two main respects. The first is in the frame of reference signified by the ring: whereas the telescope, tunnel, and funnel are all based on radial and concentric strokes within a *Polar* reference frame, the tilting carousel operates through orthogonal strokes within a *Cartesian* reference frame. The second difference is in the integration of the ring and the associated representation: in the telescope interface the ring is *overlaid*, superimposed on top of the representation; in the tunnel and funnel interfaces the ring is *inlaid*, forming a persistent part of the representational structure; while in the carousel interface the ring is *underlaid*, selectively concealed by magnified media elements. By tabulating these dimensions against one another (Figure 5), we identified three additional, distinctive focal ring designs.

<sup>2</sup> Our individual templates continued to be used throughout the full series of 15 programs, however.

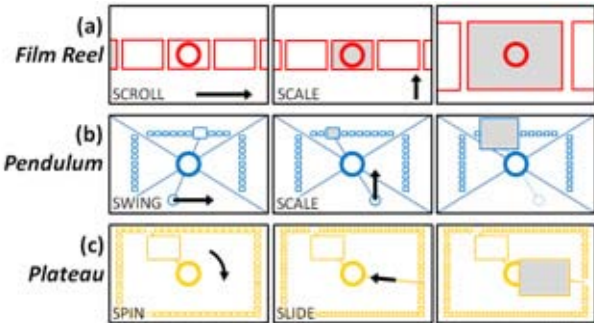


	Overlaid	Inlaid	Underlaid
Cartesian	<i>Film Reel</i>	<i>Pendulum</i>	Carousel
Polar	Telescope	Tunnel, Funnel	<i>Plateau</i>

**Figure 5. Focal ring metaphors: ring-representation integration versus frame of reference (*new designs italicised*).**

### Additional Ring Designs

Figure 6 shows the three focal ring designs suggested by the mapping of the design space (Figure 5). These are based on the three new physical metaphors of *film reel*, *pendulum*, and *plateau* respectively.



**Figure 6. Further focal ring designs.**

#### Focal Ring “Film Reel”

A focal ring could act as an index within a media collection represented as a linear film reel (Figure 6a). Stroking horizontally could scroll the film reel side-to-side, while stroking upwards could scale the film reel such that the indexed item filled the display. At this point, the ring could either fade to allow unobscured viewing of that media item, or remain visible to fulfil some further function (like acting as the origin of a marking menu).

This *film reel* design is similar to the video navigation application of Oblong Industries’ “g-speak” spatial operating system [11]. In this application, however, control is through three dimensional hand gestures at a distance from the display surface, rather than through direct touch.

#### Focal Ring “Pendulum”

A focal ring could also act as a pivot for a pendulum-like media selector (Figure 6b). Linear strokes along one edge of the display could be seen to rotate a pendulum about this central pivot, in such a way as to browse media arranged in a line on the opposite side of the ring. Moving the finger towards the line of media being browsed could then magnify the view of the selected item. Such a pendulum-based approach to media browsing avoids the “fat finger problem” of touch interaction by creating a mirrored indirection between actions and their effects.

This *pendulum* design is unlike any other interface we are aware of, although its mirrored indirection can be seen as related to the translation-based indirection of the “Shift” technique for operating pen interfaces using touch [18].

#### Focal Ring “Plateau”

Another way in which to avoid the fat finger problem is to specify a bearing from a focal ring which then

intersects a media item at the edge of the screen (Figure 6c). An item can thus be selected by spinning the finger around the focal ring, and subsequently scaled up by sliding the finger along that bearing towards the focal ring (as if drawing the item upwards towards a plateau). In such a way, multiple media items could be enlarged at the same time, with the last-touched item automatically rising to the top of the “pile”.

This *plateau* design, of having a periphery of media thumbnails and a focal area of maximised media items, shares a number of similarities with the “scalable fabric” interface for window management [14].

### ANALYSIS

The focal ring designs presented thus far support a *directed* style of manipulation, fundamentally different to the more established techniques of direct manipulation and gesturing.

In direct manipulation, the user can either manipulate the representation directly (e.g. panning a map by dragging it), or directly manipulate a peripheral control to achieve an indirect effect on the focal representation (e.g. dragging a slider to adjust the zoom level at which a map is displayed). The direct manipulation of representations relies on *explicit temporal modes* – the pointer can only act as a one tool at a time, and these tools must be explicitly switched between. Such mode switching is typically achieved through the introduction of on-screen icons, menus, or palettes, which consume potentially valuable display space. It is also possible to accidentally select the wrong tool or misremember the current selection. In contrast, the direct manipulation of controls relies on *explicit spatial modes* – the effects of actions are interpreted differently in different areas of the screen, according to the function of the underlying control. Such controls also consume display space, and require the user to divide her attention between her manipulation of the peripheral control and its effect on the focal representation.

An alternative to direct manipulation is interaction through gesturing, in which the frame of reference for interaction is typically relative to wherever the user initiates the gesture (e.g. drawing spirals anywhere to scroll within a document [10,16]). Although such gesturing allows actions to be performed directly on top of a representation, the representation itself does not provide clues as to what gestures are interpreted by the system.

Our designs differ from regular gestural controls in that the focal rings visually signify a particular frame of reference for stroking gestures directed with respect to both the ring and its associated representation. At the same time, they differ from direct manipulation in that the two dimensions of the plane can be mapped to different effects, achieving *implicit gestural modes* that reduce the need for explicit mode switching. In the following sections, we now present further analysis of these focal ring characteristics.

**Interactive Qualities of Focal Rings**

Our seven focal ring designs allow us to examine further the qualities of directed manipulation. Rather than highlighting differences in terms of ring–representation integration and reference frame, we can look for commonalities across all designs.

The greatest shared characteristic is that all interfaces combine an element of the *perceptual coupling* associated with direct manipulation (e.g. if I stroke across the display in the direction of the focal ring, the map moves in that same direction as if being dragged by my finger), with an element of the *expressive power* associated with gesture (e.g. if I stroke around the focal ring, the map will scale from the centre of the ring). The directedness of interactions also supports a degree of *control indirection*, where the focus of the user’s attention can be in a different region of the screen to their physical locus of control. This indirection can be out-to-in (e.g. with a *telescope*, stroking around the edge of the display to zoom into the centre), side-to-side (e.g. with a *pendulum*, stroking on one side of the ring to browse media on the other side), or in-to-out (e.g. with a *funnel*, crossing the central ring to change the information on the periphery). This analysis is shown in Figure 7.

This indirection of control is a valuable interaction quality, since it allows multiple effects to be chained together in the context of a single stroke, while reducing the degree to which the hands obscure these effects. Examples include:

1. In the *telescope* and *plateau* designs, a circular stroke can smoothly turn into a drag to centre, with the selected target being “lassoed” toward the ring.
2. In the *carousel*, *film reel* and *pendulum* designs, a “square wave” stroke can support many select and scale operations within a continuous gesture.
3. In the *funnel* design, the user can keep crossing over the ring to navigate through multiple levels of a hierarchy

We call these *cascade* effects. This term is both evocative of the way in which a chain of effects can “stream out of” a single continuous stroke (or multiple strokes that smoothly flow into one another), as well as being a mnemonic (CASCADE) that denotes the four significant ways in which a focal ring can be addressed through touch interactions:

- Cut Across,
- Square or
- Circle Around, or
- Directly Engage

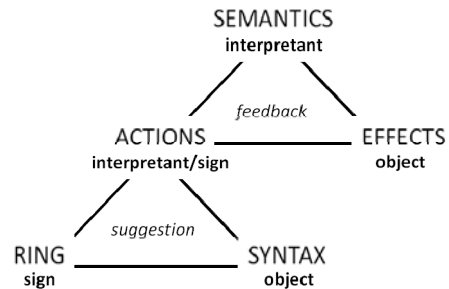
To “cut across” and “circle around” is to use a Polar reference frame, to “square around” is to use a Cartesian reference frame, while to “directly engage” is the use the area inside the ring as an *explicit spatial mode* (or “modal space” [6]) within (or from) which actions are interpreted differently (e.g. free panning within the ring in the *telescope* design for map navigation).

Physical Metaphor	Perceptual Coupling	Expressive Power	Control Indirection
Telescope	Target	Twist	Out → In
Tunnel	Slide	Stretch	Side → Side
Funnel	Spin	Cross	In → Out
Carousel	Rotate	Tilt	Out → In
Film Reel	Scroll	Scale	Out → In
Pendulum	Swing	Scale	Side → Side
Plateau	Spin	Slide	In → Out

**Figure 7. Common characteristics of focal ring designs. Operation labels taken from Figure 4 and Figure 6.**

**Semiotic Qualities of Focal Rings**

In addition to the styles of interaction supported by directed manipulation, we are interested in the question of how users come to understand focal rings as meaningful components of user interfaces. We approach this question through reference to *semiotics* – the science of signs [13]. A semiotic analysis of focal rings is shown in Figure 8.



**Figure 8. Semiotic analysis of focal rings.**

A focal ring acts as a *sign* in the context of its associated representation (e.g. a map) – it stands for something other than itself. The thing the ring stands for – its *object* – is the syntax defined by the frame of reference the ring introduces to the interface (e.g. a Polar frame of reference). The *interpretant* of this signification – the suggestion it creates in the mind of the user – is of possible actions to take with respect to the visual ring and it implied syntax (e.g. stroke toward the ring). In turn, the mental simulation of any such action can also be viewed as a sign, standing for some kind of effect on the associated representation (e.g. pan the map). The imagined feedback between actions and effects can then lead to the development of a mental model of the potential semantics of the interface (e.g. stroking towards, away from, or through the focal ring pans the map).

In such a way, even in the absence of any interaction, a focal ring can be seen to guide the user through a *chain of signification*. In the presence of interaction or personal experience with such interfaces, however, the *mode of signification* may change. To a new user, the ring may be interpreted as an unknown *symbol* – something that stands for something else through arbitrary law or convention. To an experienced user, on the other hand, a metaphor (such as “the ring is a telescope/funnel/carousel/etc.”) may help the ring to be

more readily understood as an *icon* – something that stands for something else through inherent similarity or resemblance. These *suggestions* could be confirmed or contradicted by dynamic *feedback* resulting from real operation of the interface (Figure 8). This interactive relationship between action and effect can be understood as an *index* – something that stands for something else through causation. The degree to which this causation is understood to be literally iconic (i.e. *perceptual coupling* between actions and effects), versus metaphorically iconic (i.e. *expressive power* from the application of metaphor<sup>3</sup>), versus simply symbolic, may also vary with user experience and with the strength of the underlying metaphor.

**IMPLEMENTATION**

We selected three of our seven designs for implementation, based on an analysis of their defining characteristics, their generality, and their appropriateness for their intended activity.

Of these seven designs, the *tunnel* and the *funnel* are the only ones based on abstract representations built around an inlaid focal ring. Of these two abstract representations, the *funnel* is the more general design, applicable to any tree-structure. In contrast, the *tunnel* can be seen as a specific application of the *telescope* to an abstract, tunnel-based representation of time. The *funnel* and *telescope* were therefore the first two designs selected for implementation.

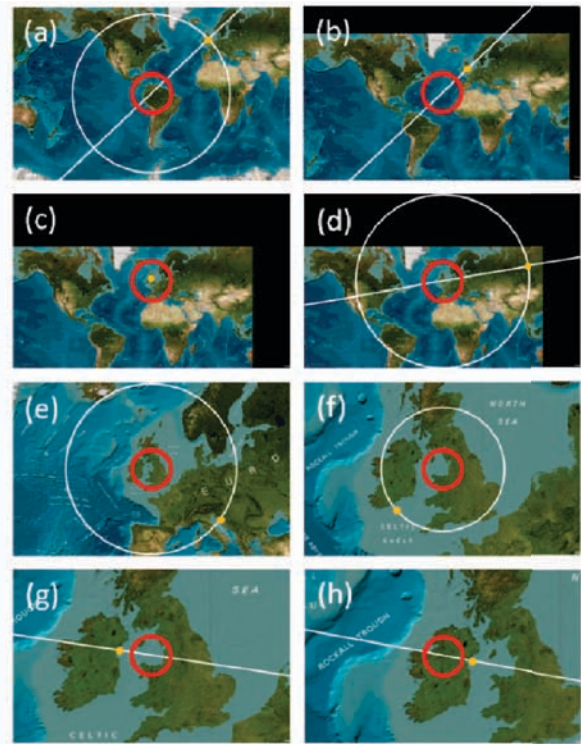
All but one of our designs (the *funnel*) incorporate a mode for *selecting* something and a mode for *scaling* that selection. All but one of these (the *telescope*) is designed to support the selection and scaling of items within a media collection. Of the *carousel*, *film reel*, *pendulum*, and *plateau*, only the *carousel* and *film reel* provide a Cartesian reference frame to complement the Polar frames of the already selected *funnel* and *telescope* designs. From these two designs, the *carousel* was chosen over the *film reel* because it had a stronger physical metaphor and made better use of the available display space.

The following section details the implementation of these three selected focal ring designs.

**“Telescope” Interface for Map Exploration**

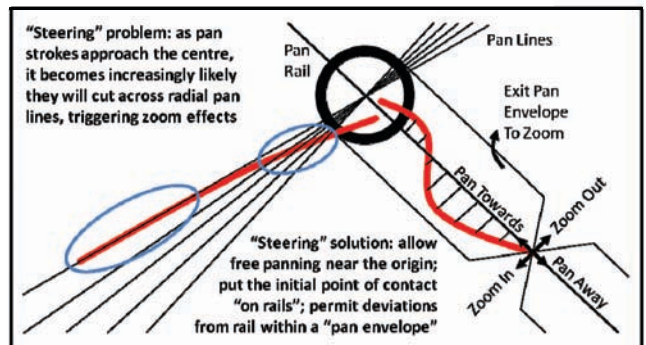
The implementation of our *telescope* design for map exploration is shown in Figure 9.

At the level of *suggestion* (Figure 8), the focal ring in this interface is like the cross-section of a stationary telescope pointing down onto the map. Locations on the map are “brought into view” by dragging them into the ring. The structure of the interface matches the structure of the task



**Figure 9.** “Telescope” interface for map exploration (*red ring and white guides emphasized from their usual transparency; orange contact point added for clarity*): (a) touch the map over the UK; (b) drag the UK towards the ring; (c) accurately position the UK in the centre of the ring by making fine adjustments in any direction; (d,e,f): zoom into the UK by circling clockwise around the ring (circling anticlockwise would zoom out); (g) smoothly change stroke direction from around the focal ring to towards it; (h) “lasso” Ireland into the centre by pulling across the ring.

– the focus of interaction (i.e. the target to zoom in on) is aligned with the focal area of the interface (i.e. the centre of the ring). Panning is naturally well supported by this syntax, since radial lines passing through the centre of the ring represent the longest possible lines along which to stroke. This syntax is not appropriate within the ring, however, since all such pan lines converge in the centre of the screen. In practice it is almost impossible not to cross from one such line to another, triggering unwanted zoom effects. This “steering” problem is illustrated in Figure 10.



**Figure 10.** Addressing the “steering” problem when using a Polar frame of reference to pan and zoom.

<sup>3</sup> For example, the *film reel* design chains the metaphors UP IS MORE and MORE IS BIG such that upwards strokes increase the size of items, and vice versa for downwards strokes.



We solve this steering problem in two ways. Firstly, we allow “free panning” within the ring: whichever way the finger moves, the map follows. Secondly, our telescope interface puts the initial point of touch contact “on rails” through the centre of the ring, such that pan strokes can deviate up to a ring radius away from their “rail” while still driving the initial contact point along it (Figure 10, right). These combined measures mean that the initial point of contact doesn’t need to be accurate, nor does the following drag to centre – the “free panning” can correct misplaced initial points of contact, and the “pan rails” are tolerant of minor deviations along the way.

The second manifestation of the telescope metaphor in this design is through the notion that “circling” around the ring can be likened to “twisting” a screw-threaded telescope, resulting in a symmetrically reversible means of zooming in to and out of the underlying map.

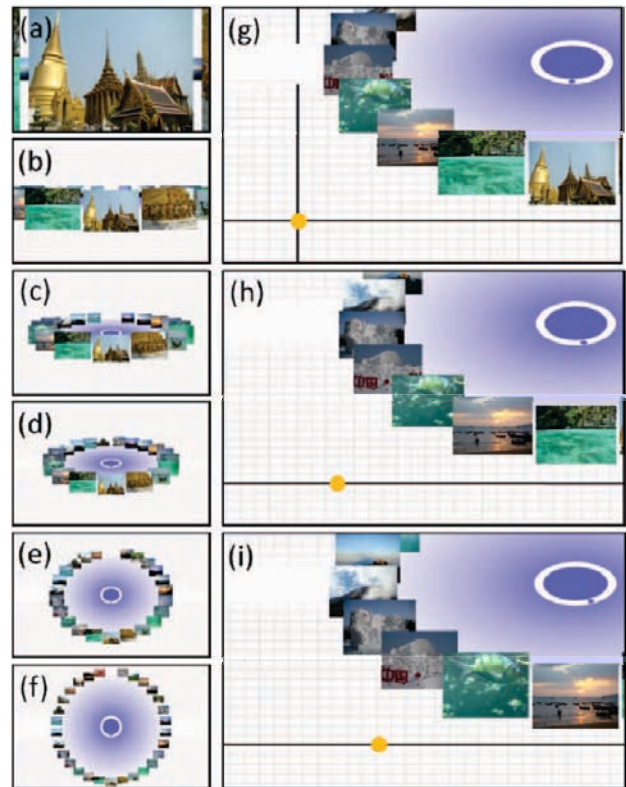
At the level of *feedback* (Figure 8), user acquisition of this interaction model, as well as correct action performance, is supported through the use of dynamic guides that give an additional level of feedback to the user. While the *telescope* interface shown in Figure 1 uses no such guides, Figure 9 illustrates how such guides can be integrated into the interface.

We can imagine showing the ideal continuations of a touch gesture on contact with the interface. Such a *contact guide* shows the radial “pan rail” and concentric “zoom circle” emanating from the point of contact (Figures 9a & 9d). Continuing to show both of these lines on commencement of a stroke constitutes a *modes guide* – potential choices for continuation of the gesture in the current or alternative mode. Finally, showing only the line corresponding to the current mode (panning or zooming) creates a *mode guide* for interaction (pan rails in Figure 9b, zoom circles in Figures 9e & 9f).

#### “Carousel” Interface for Photo Browsing

The implementation of our *carousel* design for photo browsing is shown in Figure 11.

At the level of *suggestion* (Figure 8), the focal ring in this interface is like the axis of a carousel with two degrees of freedom. The rotation of the carousel about its primary axis (perpendicular to the plane of the ring and through its centre) is achieved through horizontal strokes that are resolved into a turning moment about this primary axis. The secondary axis of the ring (lying in the plane of the ring and running horizontally through its centre) supports a “tilting” of the carousel through vertical strokes that simultaneously change the viewing distance, giving the effect of smoothly transitioning between a bird’s-eye overview of many items, and a full-screen view of the single item at the base of the carousel. Since the carousel appears to be floating in space, the background can also be utilised as a *frame guide* that signifies the appropriate actions to manipulate the carousel. In the case of the Cartesian reference frame used here, the frame guide takes the form of a grid of crossing horizontal and vertical lines (Figure 11). In the horizontal band across the screen spanning the height of the focal ring, this grid can be



**Figure 11.** “Carousel” interface for photo browsing: (a-f) tilt the carousel towards the bird’s-eye view by stroking downwards anywhere on the screen; (g) touch the screen to see a “contact guide” of potential gesture continuations; (h,i) stroke horizontally to the right below the ring to rotate the carousel anticlockwise (guides darkened and orange contact points added for clarity).

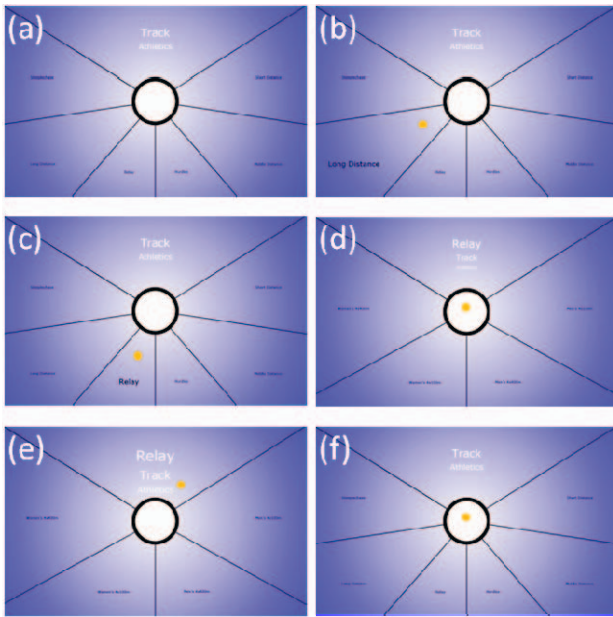
faded out to indicate the area of instability arising from almost identical horizontal strokes being resolved to opposite turning moments when performed just above or just below the carousel’s tilting axis.

At the level of *feedback* (Figure 8), the *carousel* can also be augmented with a *contact guide* (Figure 11g), as well as either a *modes guide* (like Figure 11g) or a *mode guide* (Figures 11h & 11i) in the same fashion as the *telescope*. Additionally, a dot located within the band of the focal ring can serve as an index to the start of the collection (Figures 11g, 11h & 11i).

#### “Funnel” Interface for Tree Navigation

The implementation of our *funnel* design for tree navigation is shown in Figure 12.

At the level of *suggestion* (Figure 8), the focal ring in this interface is like the neck of a funnel. Rotating the finger around the ring has no effect other than to select different sectors. The sector spanning the top edge of the screen is special – it represents the parent node of the information hierarchy. All other sectors represent children of that node. The distinction between parent and children is further indicated by differences in font – the “parent” font is large and white, whereas the “children” font is smaller and black. All sectors are triangular, pointing towards the focal ring; as such, the structure of the interface itself can be seen to suggest the intended interface action: stroke from inside a sector to inside the ring.



**Figure 12.** “Funnel” interface for tree navigation (*orange contact points added for clarity*): (a) initial view of “Track Athletics” event types; (b) touch down within the “Long Distance” sector to highlight it; (c) spin around to “Relay”; (d) cross into ring to view “Relay” events; (e) exit ring into the sector of parents (“Relay Track Athletics”); (f) cross back into ring from parent sector to return up to “Track Athletics”.

At the level of *feedback* (Figure 8), we introduced a general *progress guide* that gives visual reinforcement of the currently selected sector and the progress the user’s stroke is making towards the ring. The font of the sector name grows in size on contact within that sector, and continues to grow as that point of contact moves towards the ring. The structure of the interface can be likened to a “flattened funnel”. When one of the fanned-out child sectors is “pushed” into the neck of the funnel (the focal ring), it then “falls” onto a stack of antecedent parent nodes in the base of the funnel – the parent sector at the top of the screen. Note that this is the inverse of our earlier design – the upwards stacking of parent node labels better matches our metaphor of nodes being pushed down the neck of a funnel.

In our implementation, navigation up the hierarchy is also possible in two ways: either by stroking into the ring from the parent sector, or “pulling” out of the ring to anywhere. Whereas the former option has the advantage of permitting a *cascade* of multiple “up” navigations within a single spiralling stroke, the upwards “pull” option seems like a more natural inverse to the downwards “push” operation.

## EVALUATION

Having implemented these focal ring interfaces for map exploration, photo browsing, and tree navigation, we conducted a user study to gather preliminary data on the suitability of each interface for its activity and the feasibility of directed manipulation as an interaction style.

Each session began with a pen and paper exercise. Participants were given a piece of paper with a ring printed in the centre, and told to imagine it was a touch

screen that could detect a single point of contact. They were then asked to imagine different ways in which they could interact with the interface in relation to the ring, using a pen to mark possible finger-based interaction schemes on the paper.

Participants were then presented with each interface in turn. Initially this was in the form of a printed screenshot, on which they completed the same paper-based exercise. This was to examine the influence of the representation (map, carousel, tree), in conjunction with the focal ring, on the perceived suggestion of action possibilities. Next, they were shown each prototype, and asked to discover the appropriate style of interaction for the performance of interface-specific activities (e.g. zoom in on a location on the map, browse to a particular photo, or explore the Olympic events in a particular branch of the tree). We adopted a think-aloud protocol for this process, encouraging participants to speak their thoughts “out loud” during their exploratory interactions.

After participants had discovered or been shown the various operations of all interfaces, and consolidated this with further free exploration, they were asked to complete a questionnaire comprising both open-ended questions and questions posed as a seven-point Likert-scale.

We recruited 11 participants in total, all but 2 with computer science related majors. Each user study lasted for approximately one hour, during which time all interactions were screen captured and all speech audio recorded.

## Results

### *Pen and Paper Interaction Sketching*

The results of the initial paper-based interaction sketching suggest that a lone focal ring does not suggest the kinds of implicit frames of reference that we use in our designs. All participants chose to *directly engage* that ring as a more conventional kind of control, from rotating the ring and dragging it around the interface, to resizing the ring itself as a kind of targeting device or window onto some underlying content. In the presence of an underlying map, only three participants thought rotation was a sensible option, since “rotating a map doesn’t make sense”. More than half of the participants thought that the ring should act as a kind of focus, either dragging map locations to it (correct), or dragging it to map locations (incorrect). Only one participant guessed that rotating around the ring could be used to zoom, compared to four participants who thought resizing the ring could be used as a way of scaling the map.

For the carousel (shown in bird’s-eye view), five participants wanted to drag photos into the ring to make them larger. While this would have been correct using our *plateau* design, this was incorrect here. Three further participants wanted to tap on photos and have them automatically move and expand to fill the central region of the screen. The presence of the dot in the ring also seemed to suggest direct engagement with the ring and the dot, whereas the circular nature of the carousel representation suggested to almost all participants that



rotation of or outside the ring would rotate the photographs.

Finally, for the funnel interface only two participants considering dragging into the focal ring – everyone else thought that tapping or double tapping would let them navigate the hierarchy.

These results show that the presence of a centrally-located ring in an interface does not by itself suggest a frame of reference for interaction. It does seem to suggest a focus for interaction, but often this is not the same kind of focus actually provided by the interface. Participants' expectations were strongly influenced by their prior experience and familiarity with more conventional means of interaction, and this is unsurprising. As hypothesized earlier in our discussion of the semiotics of focal rings, participants' initial reactions were to treat the ring as an abstract symbol. How their responses changed when interacting with the actual interfaces will now be described.

*Interface Explorations*

All participants approached their experimentation with the map "telescope" interface as if it were a conventional free-panning interface. Within the first two minutes of use, only one participant fully derived the correct scheme of interaction, even in the presence of guides showing the appropriate continuations of their stroking gestures. When questioned about the role of the guides, most participants responded that they were simply a kind of "feedback" to confirm that the system was tracking their gestures. When shown the correct scheme of interaction, everyone could immediately pan and zoom the map without any difficulty. When asked whether or not the guides would be useful once the interaction model was known, no-one thought that the guides were necessary. Our forgiving "pan rails" (Figure 10) and the natural tolerance of "circle zooming" to changes of circle radius (see Figure 9) meant that these guides at best replicated feedback from the map itself, and at worst were a distraction.

In terms of interaction with the carousel, its guides elicited a similar response to those of the map. The method of operation through orthogonal strokes was more readily discoverable, obtained by all but one user within two minutes, and so these guides were reported to provide little value. However, users were still careful to stroke on top of carousel items in the style of direct manipulation. When shown that direct manipulation wasn't necessary, and that orthogonal strokes anywhere could operate the carousel, all users could immediately replicate this functionality.

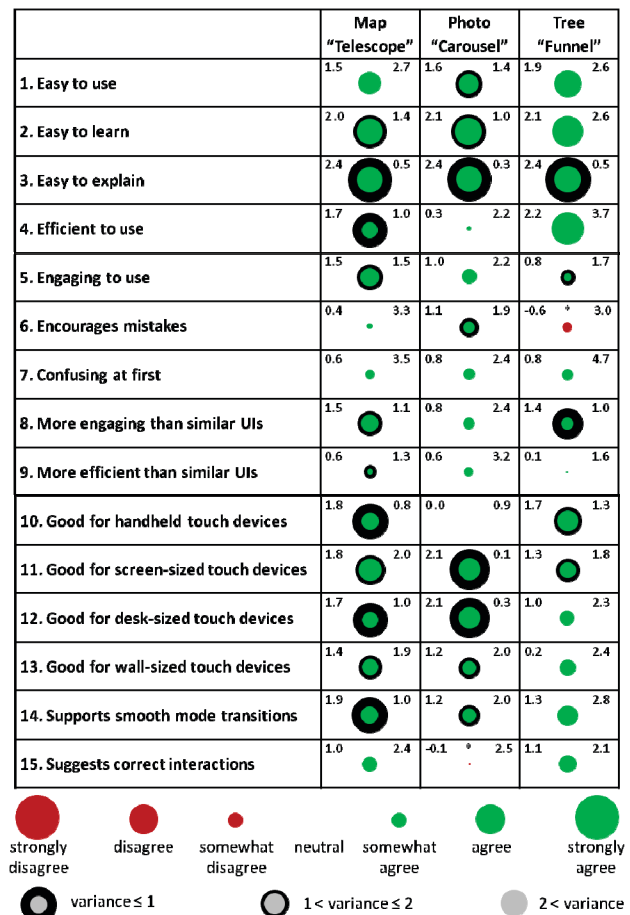
The structure of the funnel interface gave the most useful suggestion, encouraging almost all users to drag into the ring after tapping to select failed to have any effect. The "guide" of visually manipulating font sizes was found to be helpful by all participants, who could also all use font colour to distinguish between parents and children within the hierarchy. Contrary to our expectations, most participants preferred stroking into the ring from the

parent sector as the way of navigating up the hierarchy, rather than dragging out from inside the ring.

Overall, while our metaphors were useful for talking about the structure of the representation, this benefit did not necessarily extend to the suggestion of appropriate actions. However, our results imply that even though focal rings do not *inherently suggest* particular frames of reference for interaction, they do suggest *something*. Once users had associated that visual suggestion with experience of successfully operating the interface, the ring could be *understood to stand for* a frame of reference that governs the effects of directed manipulations.

*Questionnaire Responses*

Figure 13 shows a summary of average responses to the seven-point Likert scale questionnaire items. After brief periods of exposure to each of the three interfaces, participants thought that all were easy to learn, easy to use, and easy to explain to others. While the telescope and funnel designs were also thought to be efficient, the carousel was not seen as an efficient way to browse photographs. This was largely due to a reported "unnaturalness" in the mapping from finger strokes to



**Figure 13. Questionnaire responses on a seven-point Likert scale ([-3,3], 0 as neutral). Bubble size and edge thickness correspond to average score and variation respectively; red bubbles (additionally marked with an \*) represent an average response in the range of disagreement.**

carousel rotation that often led to “off by one” errors in the rotation of photos to the base of the carousel. The three designs were also thought to scale well across different display sizes, except for the carousel running on a handheld device (due to the same small usability problem, which could be remedied with further design iteration and testing). Our target interaction quality of supporting *smooth mode transitions* was also experienced to a reasonable level by these novice users, and we believe that this would increase with further exposure to such novel, focal ring interfaces.

## CONCLUSION

In this paper we have presented the concept of a *focal ring* – a graphical element that occupies a central location within a display space and defines a new frame of reference for *directed manipulation* with respect to itself and its associated representation. By assigning different functions to the two spatial dimensions of the display surface, whether in Polar or Cartesian form, we can support two *implicit gestural modes* that can be fluidly switched between in the context of a continuous, single-touch stroke. This both reduces the need for explicit mode switching and opens up new possibilities for multi-touch interactions that go beyond conventional direct manipulation.

## ACKNOWLEDGMENTS

Thanks go to Min Wang and others at MSRA for their work on the Olympic project and to Chiqing Zhang for the map control. We are grateful to the participants in our study.

## REFERENCES

1. Accot, J. & Zhai, S. (2002). More than dotting the i's – foundations for crossing-based interfaces. CHI'02, 73–80.
2. Apitz, G. & Guimbretière, F. (2005). CrossY: a crossing-based drawing application. SIGGRAPH'05.
3. Appert, C. & Fekete, J. (2006). OrthoZoom scroller: 1D multi-scale navigation. CHI'06, 21–30.
4. Arthur, K. W., Matic, N. & Ausbeck, P. (2008). Evaluating touch gestures for scrolling on notebook computers. CHI '08, 2943–2948.
5. Dachsel, R., Frisch, M., Weiland, M. (2008). FacetZoom: a continuous multi-scale widget for navigating hierarchical metadata. CHI'08, 1353–1356.
6. Everitt, K., Shen, C., Ryall, K. & Forlines, C. (2005). Modal spaces: spatial multiplexing to mediate direct-touch input on large displays. CHI'05, 1359–1362.
7. Guimbretière, F. & Winograd, T. (2000). FlowMenu: combining command, text, and data entry. UIST'00, 213–216.
8. Igarashi, T. & Hinckley, K. (2000). Speed-dependent automatic zooming for browsing large documents. UIST'00, 139–148.
9. Mitchell, K. & Kennedy, J. (1997). The Perspective Tunnel: An Inside View on Smoothly Integrating Detail and Context. 8<sup>th</sup> EG Workshop on ViSC.
10. Moscovich, T. & Hughes, J.F. (2004). Navigating documents with the virtual scroll ring. UIST'04, 57–60.
11. <http://oblong.com/> Video navigation in Oblong Industries' “g-speak” spatial operating system.
12. Olwal, A., Feiner, S. & Heyman, S. (2008). Rubbing and tapping for precise and rapid selection on touch-screen displays. CHI'08, 295–304.
13. Peirce, C. (1931–1958). Collected Papers of Charles Sanders Peirce, 8 vols. Harvard University Press, MA
14. Robertson, G., Horvitz, E., Czerwinski, M., Baudisch, P., Hutchings, D.R., Meyers, B., Robbins, D. & Smith, G. (2004). Scalable Fabric: flexible task management. AVI'04, 85–89.
15. Shneiderman, B. (1983). Direct manipulation: a step beyond programming languages. IEEE Computer 16, 8 (August 1983), 57–69.
16. Smith, G.M. & Schraefel, M.C. (2004). The radial scroll tool: scrolling support for stylus- or touch-based document navigation. UIST'04, 53–56.
17. Smith, R.B. & Taivalsaari, A. (1999). Generalized and stationary scrolling. UIST'99, 1–9.
18. Vogel, D. & Baudisch, P. (2007). Shift: A Technique for Operating Pen-Based Interfaces Using Touch. CHI'07
19. Wang, S., Poturalski, M., Vronay, D. (2005). Designing a generalized 3D carousel view. CHI'05, 2017–2020.