D-MO: Depth from Motion and Occlusion as a Visual Channel for Information Visualization

Carla Coutant, Adrien Chaffangeon Caillet, Renaud Blanch

D-MO: Depth from Motion and Occlusion as a Visual Channel for Information Visualization

Carla Coutant
coutantc@univ-grenoble-alpes.fr
Univ. Grenoble Alpes, CNRS,
Grenoble INP, LIG
Grenoble, France

Adrien Chaffangeon Caillet
chaffana@univ-grenoble-alpes.fr
Univ. Grenoble Alpes, CNRS,
Grenoble INP, LIG
Grenoble, France

Renaud Blanch
blanch@imag.fr
Univ. Grenoble Alpes, CNRS,
Grenoble INP, LIG
Grenoble, France

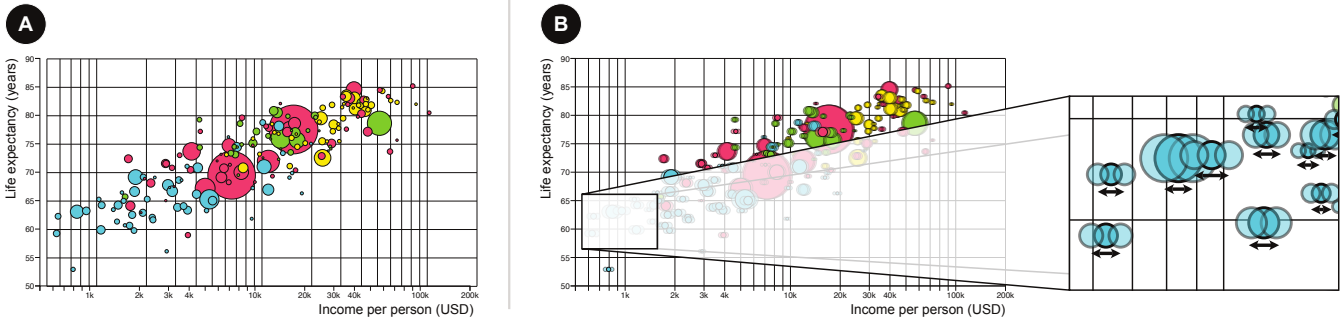

Figure 1: Visualization of development indicators by country: (A) Gapminder's original version [15]; (B), with an additional attribute encoded using D-MO (shown with transparency at the extreme positions of the oscillations).

Abstract

On a visualization, the position of the marks encoding data is the most expressive and effective visual channel. It conveys order and quantity without impairing the perception of other visual channels. In the field of Information Visualization, position is often restricted to two dimensions, because using the third dimension, *depth*, usually affects the perception of size, which is also one of the most effective visual channels.

We propose a new visual channel, D-MO (Depth from Motion and Occlusion), a combination of visual cues, *motion* and *occlusion*, with interactions, that induces a depth perception suitable for combined use with classic visual channels. We characterize the expressiveness and effectiveness of D-MO and show that it is a magnitude channel with good accuracy, acceptable discriminability, and is separable from size. Thus, D-MO opens up new areas for visualization design, which is limited by the scarceness of available visual channels.

CCS Concepts

• **Human-centered computing** → **Information visualization**; **Visualization techniques**; **Visualization theory, concepts and paradigms**; **Empirical studies in visualization**.

Keywords

Depth, information visualization, visual channel.

CHI '26, Barcelona, Spain

ACM Reference Format:
Carla Coutant, Adrien Chaffangeon Caillet, and Renaud Blanch. 2026. D-MO: Depth from Motion and Occlusion as a Visual Channel for Information Visualization. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26), April 13–17, 2026, Barcelona, Spain.* ACM, New York, NY, USA, 18 pages. https://doi.org/10.1145/3772318.3791042

1 Introduction

To create a visualization, a sequence of steps (known as the information visualization pipeline [9, 12]) transforms data into a visual representation. A crucial step in this process is visual encoding, which assigns visual channels (e.g., the position, size, or color of marks) to data attributes[1] (e.g., life expectancy, GDP per capita, population, or world region in the case of development indicators by country, as in the Gapminder bubble chart [15], where marks are circular, see Figure 1). This encoding must be designed with particular care, as the list of available visual channels is limited. Depending on the taxonomy, the number of channels varies from eight, according to Bertin's original list [2], to thirteen or fourteen according to Mackinlay or Munzner [29, 34] (though these supplemental channels cannot all be used simultaneously). These visual channels do not share the same characteristics. They can convey different kinds of information and are suited to different types of data attributes (nominal, ordinal, etc.) This ability to convey information based on the nature of attributes is characterized by a set of properties. According to Bertin [2], those properties are: *association*, *selection*, *order*, and *quantity*; while Munzner [34] proposes to use *expressiveness* and *effectiveness* (the latter including *accuracy*, *discriminability*, *separability*, *grouping*, and *emergence*).

[1] We use Munzner's terminology [34]: *visual channels* for color, size, etc., and *data attributes* for the columns of data tables.

Those properties allow for the classification and ranking of visual channels. All taxonomies agree on the most expressive and effective ones, namely, the horizontal and vertical dimensions of position, followed by the (1D or 2D) size of the marks used to encode the data items[2]. Depth comes after 2D size in Munzner's list, and its use is discouraged, as we see below.

When the data is not inherently three-dimensional, the use of the depth dimension of position is generally discouraged in the Information Visualization community[3]. There are many reasons for this recommendation, the most obvious being that depth is typically represented using a perspective projection (whether on a standard screen or a stereoscopic display), which affects the apparent size of elements. The perceived size of marks in a visualization then becomes a combination of two data attributes: the one encoded by size and the one encoded by depth. Then, we may wonder when looking at a given mark, if it appears small because its size encodes a small value, or because its depth encodes a large value. In this case, introducing depth as a visual channel comes at the cost of ambiguity in size which is one of the few visual channels effective at conveying quantity [13, 23, 34].

In this paper, we present the design and an initial evaluation of a new depth visual channel: D-MO, which introduces a Depth dimension based on Motion —more specifically, *structure from motion*— and Occlusion. D-MO is inspired by *wiggle stereoscopy*, which gives "an illusion of 3D by showing two images in rapid alternation" [18]. D-MO does not rely on perspective, in order to avoid interference with the size channel, and it uses an animation between the angles of view to make the transition between them smooth. Figure 1B, tries to convey this animation on the right: in this static illustration, the positions of the marks over time are superimposed with transparency. D-MO also allows users to interactively control the angle of view if desired.

To characterize the D-MO channel, we conducted a first controlled study reported in Section 5. The task involved estimating the depth of two specified targets in bubble charts enhanced with D-MO. Our results show that D-MO is a channel that shares to some extent *expressiveness* and *effectiveness* of position and size [13, 23, 34].

We also investigated how D-MO affects size perception. To do so, we conducted a second controlled study reported in Section 6. The task involved estimating the size of two specified targets in bubble charts enhanced with D-MO. Our results show that D-MO only slightly affects the perception of size, but does so to a much smaller extent than other biases affecting size perception.

We finally show that D-MO is usable as a visual channel in real visualizations. To do so, we conducted a pre-study reported in Section 7. The task involved answering questions about a bubble chart enhanced with D-MO, which encodes the data from Gapminder [17]. Our results indicate that participants made minimal errors, with the majority of them not being associated with D-MO. To further illustrate the versatility of D-MO in visualization, we have integrated it into diverse visualizations, including bubble charts, parallel coordinates, and terrain visualizations.

[2]Tilt/angle appears before area in Munzner's list, but this channel has specific use cases as it is not compatible with all shapes.
[3]The largest section of Chapter 3 *Rules of the thumb* of Munzner's book [34] is titled *No Unjustified 3D.*

2 Related Work

Our state of the art focuses on the use of depth in Information Visualization. We discuss various approaches that explore how a three-dimensional (3D) structure can be reconstructed from motion, relying on the general framework established in psychology known as *structure from motion* [55]. This phenomenon encompasses specific approaches, such as *motion parallax*, the *kinetic depth effect*, and the *stereokinetic effect*, all of which rely on our perceptual system's ability to exploit the relative movement between visual elements. Although this relative movement is an underlying principle of all these approaches, it also constitutes an independent depth cue on its own. For this reason, the final part of this literature review focuses specifically on relative motion.

2.1 Depth in Information Visualization

Several depth cues exist. It is possible to distinguish between monocular cues, which can be perceived using only one eye, and binocular cues, which require the simultaneous use of both eyes. The latter rely on binocular disparity, which refers to the difference in position between the images formed on each retina due to the distance between the eyes, as a cue for the brain to perceive depth [64]. In the context of Information Visualization, binocular disparity is primarily leveraged through specialized equipment, such as stereoscopic displays. These enhance depth perception but are effective only for 3D data, for complex tasks, or when other depth cues are not available [30].

The vast majority of studies conducted in experimental psychology, including those mentioned in this section, have been carried out under monocular conditions [4, 5, 7, 10, 11, 14, 19, 22, 27, 36, 38, 41, 46, 47, 66–68]. Two commonly used monocular cues are occlusion and perspective. Occlusion is a strong depth cue —described by Munzner as "the most powerful" [34] and by Ware as "probably the strongest depth cue" [60]— which explains its frequent use. Furthermore, perspective is often employed in 3D visualizations. Several 2D visualizations can be adapted into 3D, such as bubble charts [40], maps [54], graphs [45] or treemaps [43]. Nevertheless, these visualizations require a navigation technique, often complex, to explore these 3D structures [34]. Moreover, since perspective affects size perception, it is mainly recommended for data with an inherent 3D structure [34]. Indeed, while our perception of 2D dimensions or 1D size is linear, depth perception under perspective is not [52], and that is why it is ranked after 2D size by Munzner. Among other depth cues [60], we can cite relative size (which also interferes with the use of size as a visual channel); shadows and color attenuation (which affect color perception); as well as motion parallax and kinetics.

2.2 Motion Parallax and Depth Perception

Gibson et al. [19] defined motion parallax as the "optical change of the visual field of an observer which results from a change of his viewing position". In particular, objects closer to the observer appear to move faster across the field of view than distant objects [56]. This velocity difference provides information that can function as a depth cue. In previous works, motion parallax was described as an effective depth cue [14, 46, 47]. It can provide rapid depth perception, occurring in approximately 30 ms as reported by Nawrot

and Stroyan [36]. Motion parallax is all the more effective when self-produced rather than externally generated [46]. Rogers and Graham [47] suggested that while self-produced parallax helps clarify the direction of depth perception, it is not the primary factor determining whether depth is perceived. They concluded that it is the spatial frequency characteristics of the velocity field, along with the temporal frequency of velocity changes, that determine whether motion or depth is perceived.

Motion parallax is accurate for relative depth judgments [46, 47, 67]. Indeed, in their second experiment, Yoonessi and Baker [67] concluded that motion parallax is effective for depth ordering and depth magnitude estimation, for small depth differences and gradual depth changes. Furthermore, Rogers and Graham [46] found that motion parallax provides quantitative depth information. However, Caudek and Proffitt [11] observed that observer-relative motion does not contribute to depth magnitude, but depth order only. They suggested that depth magnitude is explained by object-relative motion. Likewise, Gibson et al. [19] stated that motion parallax helps in separating objects into different depth layers but does not provide absolute distance information.

Rogers and Graham [46] noted that, in random-dot motion-parallax displays, perceived depth increased with the amplitude of relative motion. Nevertheless, at the largest amplitudes, the increase in depth perception was less than anticipated, suggesting that "large amplitudes of relative movement do not produce correspondingly large amounts of perceived depth".

Dynamic perspective cues (vertical displacement, lateral speed gradients and speed differences between near and far surfaces) enhance depth perception from motion parallax [7]. More specifically, differential velocity —differential motion speed of objects at different depths, corresponding to the traditional definition of motion parallax— provides a highly effective relative depth cue (except for children) as reported by Carpenter [10].

Loomis and Eby [27] suggested that circular variants of motion parallax in dot diplays can elicit coherent, non-rigid 3D shape from motion even though the pattern does not correspond to any rigid 3D motion, indicating that a physically rigid object is not required for depth perception.

Faubert [14] found that binocular motion parallax and binocular disparity each provide effective relative depth judgments on their own. When both cues are combined, performance in depth judgment increased significantly compared to using either cue alone. Moreover, it has been highlighted that motion parallax and stereopsis produce a similar depth experience [46, 47], justifying the interest of motion parallax as a depth cue.

Expansion-compression (i.e. relative motion of texture elements) has been found to be reliable for perceived depth ordering at smaller depths, while accretion-deletion (i.e. covering and uncovering of parts) enhances depth ordering at larger depths [38, 68]. Furthermore, Yoonessi and Baker [68] asserted that accretion-deletion alone does not create depth perception, even though it theoretically provides all necessary information. Accretion-deletion only works when combined with expansion-compression. This confirms that both cues complement each other, creating a stronger depth perception than either cue alone.

2.3 Kinetic Effects and Depth Perception

Another visual phenomenon inducing a depth perception is the kinetic depth effect (KDE), which was first introduced by Wallach and O'Connell [59]. The KDE refers to the phenomenon in which the deforming shadow of a three-dimensional object, rotating between a punctiform light source and a translucent screen, is perceived as a rigid three-dimensional object in rotation.

In moving-dot displays, motion alone is enough for KDE, without the need for dot density changes. However, Sperling et al. [50] observed that performance in their identification task improved with an increase in the number of dots in the display while the magnitude of depth perceived also increased.

Like the KDE, our approach leverages motion as a means to convey depth without requiring observer movement. However, instead of relying on object rotation in three-dimensional space, it utilizes two-dimensional motion patterns, like the stereo kinetic effect (SKE).

SKE occurs when a two-dimensional pattern (e.g., nested circles) rotates within the image plane, creating the illusion of a three-dimensional form (e.g., a cone or funnel) [41]. This phenomenon was first observed by Musatti in 1924 while studying stereokinetic phenomena, and later formally defined [35]. SKE can produce convincing depth with simple two-dimensional motion [41], reducing computational load [26].

Both SKE and KDE primarily rely on the structured motion of a single object or contour set, where depth emerges from coherent rotational transformations rather than from variations in motion across multiple elements. However, depth can also be conveyed through relative motion differences [66]. The following section explores how relative motion, without structured object rotation, can effectively convey depth in various contexts.

2.4 Relative Motion as a Depth Cue

It has been shown that relative motion of texture and contour at an edge can specify depth order [66], demonstrating that this is a powerful depth cue. In studies on rotating dot patterns, Braunstein [4] demonstrated that increasing the number of dots enhances depth perception, following a logarithmic trend, wherein larger increases in the number of dots lead to progressively smaller gains in perceived depth.

According to Braunstein, greater perspective enhances the perceived depth of rotating dot patterns, although its effect is weaker than the influence of numerosity. His results further showed that perspective contributes most strongly when only a few dots are present. Braunstein also reported that increased perspective reduces perceived coherence, defined as the extent to which the dots appear to maintain a rigid, stable configuration during motion.

Several studies found that suggestion could influence whether motion was seen as two-dimensional or three-dimensional [5].

2.5 Gaps in the Literature

As shown in the studies referenced above, motion-based depth cues have been extensively studied in perception research, particularly in the context of motion parallax, KDE, and SKE. However, despite their effectiveness, these traditional approaches present several limitations that our work seeks to address. This section highlights

the primary gaps in the literature and positions our approach as a novel visual channel.

Firstly, by definition, motion parallax requires observer movement to perceive depth [19], which can limit applicability in visualizations where the observer remains static. On the other hand, depth perception in binocular stereoscopic displays relies on specialized hardware (e.g., stereoscopic monitors, head-mounted displays, etc.) [14, 24, 47, 58]. This dependence on observer movement or additional equipment makes those techniques less practical for many visualization applications. In contrast, our approach does not require observer motion, nor specific equipment. Instead, we provide an alternative that leverages differences in motion properties as a depth cue, making it suitable for standard 2D displays.

As discussed earlier, KDE and SKE demonstrate that depth can emerge from the structured rotation of a single object or contour set. Our approach diverges from these traditional motion-based depth cues by encoding depth through an oblique projection whose axis oscillates from left to right. Rather than requiring a coherent rotational transformation, we explore how variations in oscillation amplitude and velocity among separate objects contribute to depth perception.

Finally, very little research has explored wiggling as a direct depth encoding mechanism [32, 37, 39]. Our work fills this gap by extending motion-based depth perception by systematically evaluating how motion differences across objects contribute to depth perception in a binocular setting.

In Information Visualization, the use of depth is generally discouraged because perspective distortion affects the perception of other visual channels, especially planar spatial position and size [34]. Nevertheless, depth perception can be induced by other types of stimuli, such as motion. Because motion is strongly separable from static visual channels (e.g., color, spatial position), it does not interfere with their perception [34]. According to the literature, these types of motion are promising for Information Visualization, as they enable the distinction and ordering of different depth levels. In the following study, we measure the effectiveness of these motion cues for Information Visualization.

3 Methodology

Munzner proposes a framework to structure the different levels at which visualization should be validated [33, 34]. We do not focus on the lowest level, *algorithm design*, since the D-MO implementation is straightforward from the description given in Section 4. Going up towards data, we first stop at the *visual encoding and interaction design* level, which is the main focus of our contribution, and detail below the methodology we have chosen for this validation done with two controlled experiments reported in Section 5 and Section 6. We then propose a pre-study reported in Section 7 to show that D-MO can be integrated into existing visualizations and allows to perform usual tasks on a dataset, a first step towards a validation at the *operation and data type abstraction* level, even if more work has to be conducted here to gain in generality. We detail below the taxonomies of tasks from which we drew inspiration. Finally, the *domain situation* validation is not addressed here since D-MO is not yet integrated into fully-fledged visualization systems.

3.1 Characterization of Visual Channels

Existing lists of visual channels are often structured and ranked according to various criteria and metrics that are not always explicitly defined. Bertin's notions of *association*, *selection*, *order*, and *quantity* [2] are defined in broad terms and no actual way of precise characterization, apart from his own expert judgement, is given by the author. Munzner's *expressiveness* and *effectiveness* [34] are more precisely defined and use some quantitative metrics to characterize visual channels. Our experiments on depth (Section 5) and size (Section 6) perception with D-MO use those metrics as explained in Section 8.1. This allows us to compare D-MO with other visual channels without running experiments with those channels as baselines, since, e.g., Stevens' exponent [52] or Cleveland and McGill log $error$ [13] are known for existing visual channels.

For the depth perception experiment, we used a magnitude estimation protocol where the estimation was reported by the participants on a rectilinear scale that acted as a psychophysical scale. The minimal and maximal values were explicitly shown during training to provide a common *constrained scaling* shared among participants [61, 62]. For the size perception experiment, we used a magnitude production protocol where participants were asked to match the size of the marks to avoid potential bias introduced by Stevens' exponent on the size perception. For both experiments, we recruited sixteen participants, following MacKenzie's advice to "use the same number of participants as in similar research" [28], and basing our number on works on the characterizations of visual channels made in the Human-Computer Interaction and Information Visualization communities (e.g., 12 participants for Widgor et al. [63], or 16 for Bezerianos and Isenberg [3]). This number is between the mode (12) and the median (18) for controlled experiments reported in CHI 2014 papers [8], which is consistent with numbers (mode and median: 18) reported for CHI 2019 experiments filtered to only consider lab studies with *visualization* appearing in their topic [42].

3.2 Typology of Tasks

Regarding task abstraction, we outline below three task taxonomies. These taxonomies do not share the same granularity and trying to unify them is out of the scope of this paper.

Firstly, Munzner [34] defines three levels of actions to define user goals: high-level goals to specify how a visualization can be used to be analyzed; mid-level goals focusing on search; and low-level goals focusing on query. The high-level actions are divided into two types: consume goals (including *discover*, *present*, and *enjoy*), and produce goals (including *annotate*, *record*, and *derive*). At a finer level, mid-level goals are divided into four alternatives, depending on whether the identity and location of the search target are known. The *lookup* alternative corresponds to cases in which the user knows both the target and its location. The *locate* alternative is used when the target is known but its location is not. In the *browse* alternative, the user knows the location but not the target. Finally, in the *explore* alternative, both the target and the location are unknown. For the low-level goals, Munzner describes the *identify* scope, which returns characteristics of known targets or specific references; the *compare* scope for comparing multiple targets; and the *summarize* scope to provide an overview.

In addition, Amar et al. [1] present ten low-level analysis tasks: *Retrieve value*, *Filter*, *Compute derived value*, *Find extremum*, *Sort*, *Determine range*, *Characterize distribution*, *Find anomalies*, *Cluster*, and *Correlate*. For synthesis reasons, we do not describe these tasks in detail, as their names are self-explanatory.

Lastly, we describe in more detail a study by Szafir et al. [53] that divides tasks into four main categories. The first, "identification tasks" —searching for specific data points or subsets— includes *absolute-value identification* (isolating points matching a specified value); *relative-value identification* (finding items relative to the distribution); and *outlier identification*. The "summary tasks" category —estimating statistical properties— includes *mean and variance estimation*; *distribution statistics* (e.g., skewness, kurtosis); and *numerosity* (estimating the quantity of a set of data points). The "segmentation tasks" category —organizing data points into subsets— contains *segmentation by spatial position* (grouping items based on spatial layout); and *segmentation by features* (grouping by featural similarity). Finally, the "structure estimation tasks" category —extracting patterns not captured by statistics— includes *trend detection* (identifying relationships between variables); *similarity detection* (comparing patterns across regions); and *distribution shape* (recognizing the overall data shape).

4 Depth from Motion and Occlusion (D-MO)

D-MO combines motion and occlusion cues to convey depth perception. To achieve this, we drew inspiration from *wiggle stereoscopy*, which gives "an illusion of 3D by showing two images in rapid alternation" [18]. D-MO also alternates between two points of view while adding a third intermediate one. It produces a continuous transition between those three points of view. D-MO does not use a rotation around a vertical axis, because such a rotation would induce both lateral and back and forth movements of the marks, especially near the edges (see the video figure). Instead, it uses an oblique projection —parallel to an axis that is not orthogonal to the projection plane— where the axis angle oscillates from left to right. The projection angle is interpolated over time to create a continuous animation between the left and right images of the wiggle stereogram, and a third, central position. The angle can also be directly manipulated by users through a drag interaction. The animation and interaction are described below.

Regarding occlusion, the standard drawing order in visualizations such as bubble charts involves painting marks in decreasing size, placing smaller ones in front of the larger ones to avoid occlusion as much as possible. This order, based on the data attribute encoded by size, can be different than the one encoded by depth. In a pilot study, we validated that occlusion is a strong depth cue. Therefore, it is essential to paint marks according to the order defined by the depth-encoded attribute to make D-MO effective.

D-MO can be used with two modes: with or without interaction, depending on whether the projection axis is controlled by a time-dependent function or directly by the user. To facilitate understanding of this technique, we provide an example of D-MO used on a real dataset, available online[4].

We explored different designs for the animation in a pilot experiment with few participants to settle on the design described hereafter. The participants were able to change various parameters of the movement: its direction (horizontal/vertical/combination of both), its frequency (whether it stops in the middle position), and different interpolations. Several designs for the interaction were also tested. We present below the combination that satisfied all participants and that was used in the subsequent experiments.

[4]D-MO online demo. <iihm.imag.fr/blanch/projects/d-mo>.

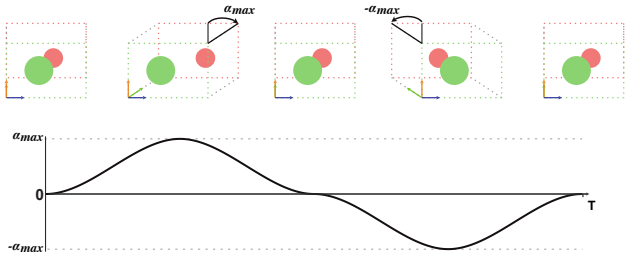

Figure 2: Temporal evolution of the projection angle α.

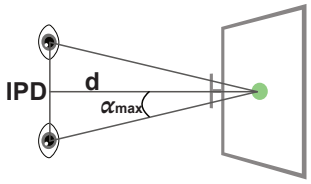

Figure 3: Maximum projection angle (α_{max}) function of screen distance (d) and interpupillary distance (IPD).

4.1 Animation

Instead of using two viewpoints, as in *wiggle stereoscopy*, we add a central position, as well as a continuous transition between these three angles. According to our pilot study, the stop at the central position marks the exact horizontal position of the marks, and therefore minimally disrupts judgments related to the attribute encoded by the X-axis. It also facilitates interaction, as it corresponds to the position that will spontaneously be targeted when the goal is to reach a specific mark.

The reference plane used for the projection has a depth of zero. Marks located in this plane remain stationary, while those positioned in the back move in the opposite direction of those in the front. The amplitude of the movement is proportional to the distance from the reference plane. The animation cycles through the central, left, central, and right angles in sequence (0, α_{max}, 0, $-\alpha_{max}$), repeating continuously. Interpolation between these angles is achieved using the *smoothstep* function, defined over the interval $[0, 1]$ as: $x \rightarrow 3x^2 - 2x^3$.

Figure 2 shows the evolution of the projection angle on a single cycle (bottom) and its effect on two planes in a simplified visualization (top). The movement amplitude, α_{max}, is set to $\pm 2.5°$, in such a way that the extrema correspond to the angular difference between the left and right eyes of a user sitting approximately 70 cm from a screen, with an average interpupillary distance of 6.3 cm [20] (Figure 3). Finally, the movement frequency is set to 0.6 Hz (3 cycles every 5 seconds), based on the results of the pilot study.

4.2 Interaction

The animation is combined with interactions that introduce a coupling between the projection and the user's actions, similarly to motion parallax.

The first interaction consists of using the depth of the mark being hovered over by the mouse as the depth of the reference plane for the projection. This mark then becomes stationary: all marks located in front oscillate in one direction, while those located behind oscillate in the opposite direction. This design replicates an attention mechanism that is present in perception, where the object of focus appears still while the foreground and background planes move. This interaction is also compatible with standard mouse-over interactions, which display details of the hovered mark or indicate its projection on the axes using cursors. To avoid instabilities caused by the movement of the marks, the active zones used for picking correspond to their central position and do not follow their motion. The pilot study confirmed that pausing the animation at this central position makes it a natural target for interaction.

The second interaction consists of directly manipulating the axis of the oblique projection through dragging. When the mouse button is pressed, the cursor's position on the reference plane is recorded. The animation is then paused, and the projection axis is set orthogonal to the projection plane (corresponding to the central position of the animation). This axis passes through the point designated by the cursor, which is considered the pivot for subsequent dragging movements. When the mouse moves, the cursor's position designates a point on a plane parallel to the reference plane, placed in front of it at a distance equal to half the amplitude of the depth axis. The projection axis is defined as the line passing through this point and the pivot established at the start of the interaction. Marks located at the depth of the projection plane remain stationary (as in the animation); those in front follow the mouse movement with an amplitude that depends linearly on their distance from the reference plane; and those behind move symmetrically to the mouse, relative to the pivot. When the mouse button is released, the axis returns to its central position and the animation resumes. This interaction replicates the motion parallax mechanism, in which the movement of objects in the visual field is coupled with the movement of the observer.

To determine the properties of D-MO, both with and without interaction, we conducted the experiments described below.

5 Characterization of Depth Perception Induced by D-MO

In this section, we present a user study conducted to characterize the properties of the D-MO visual channel, relying on established metrics used to assess visual channels' expressiveness and efficiency.

5.1 Visualization and Data

In this experiment, we chose a scenario that required encoding a quantitative data attribute, while the conventional visual channels that allow quantity assessment (position and size) were already allocated to other quantitative attributes. Depth was then used as an additional visual channel. The visualization for this dataset was a bubble chart that encoded five attributes. Four of these attributes were quantitative: two were encoded using the two standard dimensions of position (x and y), one using the size of the marks, and the last using D-MO. The fifth attribute was nominal and encoded using the color of the marks. We used filled marks, as colors are more discriminable in filled shapes than in unfilled shapes [49]. Moreover, we opted for a bubble chart visualization in order to incorporate an additional dimension, compared to a scatterplot, through the use of mark size. By using a single mark shape, we avoided biases resulting from the significant influence of shape on size perception [49].

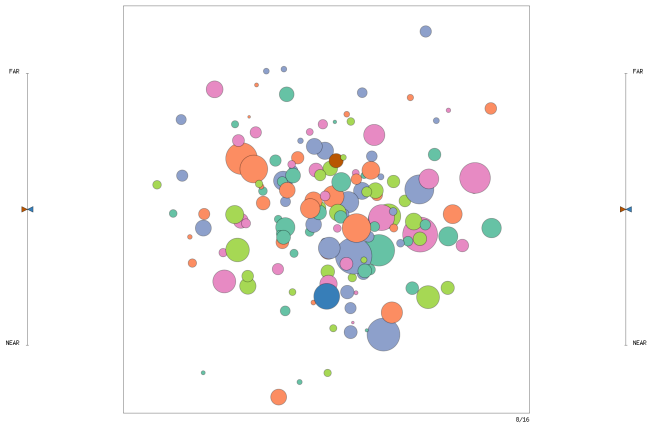

Figure 4: Task of the experiment: participants must estimate the depths of the red and blue bubbles and report them using the sliders placed on the sides of the visualization.

A random dataset was created for each task described below. The attributes encoded by the position followed normal distributions; the attribute encoded by the size followed a log-normal distribution; and the nominal attribute followed a uniform distribution over five categories. Finally, the attribute encoded by D-MO was uniformly sampled from the interval $[-1, 1]$. This sampling was not random, as we wanted to control depth to facilitate analysis. All distributions were independent; thus, there was no correlation across attributes.

The visualization was displayed within a square centered on the screen. Its sides measured three-quarters of the screen height (2160 pixels). Mark positions were scaled so that the sides of the square corresponded to three standard deviations of their normal distribution. Mark sizes were scaled so that the average of their diameters represented 2.5 % of the square's side, i.e. 54 pixels. Colors corresponded to the first five colors of the *ColorBrewer Set2* palette [6]. Figure 4 shows the screen as presented during a task.

The experiment software was implemented in Python3 using PyOpenGL[5]. The experiment was conducted on an iMac Retina 5K 27 inches (2014). The computer was placed on a desk, 70 cm from the chair where participants were seated. Participants used a keyboard and a mouse to enter their responses. The setup was installed in a relatively bright room.

5.2 Task

For each task, two targets were selected among the marks and highlighted in red and blue. Participants were asked to estimate their depths. In addition to magnitude estimation, which measures how accurately users perceive data values presented through graphical encodings [48], we used *constrained scaling* [61] to report magnitude on the vertical linear sliders that acted as psychophysical scales. This is similar to Saket et al.'s use of sliders to record participants' judgments for position encoding [48].

We placed two identical, synchronized vertical sliders on both sides of the screen (Figure 4). The lower and upper ends of each slider corresponded, respectively, to the nearest (*near*) and furthest (*far*) bubbles in the chart. The training began with targets positioned at the nearest and furthest positions on the slider. Participants were told that these positions corresponded to the slider's endpoints. Each slider featured two adjustable cursors, one for each target, indicated by their corresponding colors. Participants were free to manipulate either slider. Using the mouse, they adjusted the cursors according to the perceived depths of the red and blue targets. Each cursor could be moved as many times as needed and in any order. To validate their response and proceed to the next task, participants pressed the space bar on the keyboard.

The two targets were selected from a set of depths that subdivided the depth interval $[-1, 1]$ into 20 equal parts. To reduce the number of combinations, we selected five median depths $\{-0.8, -0.4, 0, 0.4, 0.8\}$ and five possible offsets $\{\pm 0.1, \pm 0.2, \pm 0.4, \pm 0.6, \pm 0.8\}$ and combined them. Figure 5 shows the 16 combinations of median depths and offsets (depth is represented on the vertical axis) kept for the experiment. These 16 combinations were designed to balance evenly the depth positions and satisfy the following three constraints. First, the 11 selected depths that were multiples of 0.2 (indicated by horizontal lines) were each used in exactly two combinations, while the 10 intermediate depths were used in only one. Second, the red and blue targets appeared equally often in front and in back. Last, the respective mean depth of the red and blue targets was the same and centered at zero.

To verify that the density of the marks did not affect depth perception, we alternated the number of bubbles across tasks using three values: 81, 121 and 161 (as shown in Figure 6). This approach allowed us to explore a typical number of marks in Information Visualization, even though it did not consider very large datasets.

[5]The code is available at <iihm.imag.fr/blanch/projects/d-mo>.

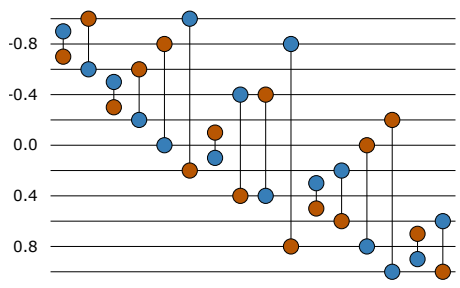

Figure 5: Depth configurations of target pairs.

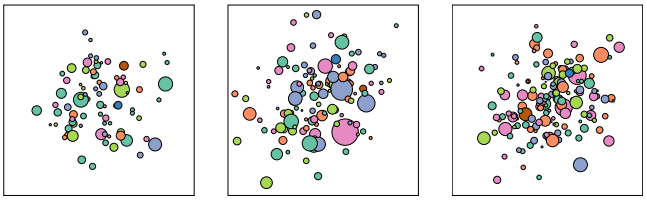

Figure 6: Density levels: 81, 121 and 161 marks.

Since the bubbles were uniformly distributed across depth, this strategy ensured that each of the 21 depth planes of the red and blue targets always contained bubbles.

Since mark positions and sizes were randomly generated, we verified that the red and blue targets satisfied three criteria:

- they had to be large enough, i.e. with a radius greater than 27 pixels (mean radius);
- their positions had to remain relatively central, i.e. within 0.75 standard deviation from the mean;
- they should not be significantly occluded by any other mark, i.e. when the projection angle was zero, a circle with a radius equal to one-quarter of the target's radius had to remain visible around its center.

These three conditions ensured that the targets were not too difficult to locate, since our focus was the evaluation of the depth perception, not the target visual search. If the three conditions were not met, another dataset was generated.

5.3 Procedure

First, participants were asked to fill out a consent form outlining the purpose of the experiment. They were then seated in front of the computer running the experiment software.

The experiment was divided into two phases: one with interaction and one without, as described in Section 4. Half of the participants began with the non-interactive phase followed by the interactive phase, while the other half followed the reverse order.

For each phase, 16 target configurations were generated for each of the three levels of mark density, resulting in 48 tasks. These tasks were then randomly distributed into three blocks of 16 tasks each. Before the three blocks of each phase, an additional training block of 16 tasks was presented to the participant. This training block consisted of tasks selected from the pool of 48 tasks corresponding to that phase. The task order within the training blocks was controlled so that it began with tasks involving extreme front and back positions, allowing participants to calibrate their depth judgments. The two training blocks began with the following three tasks: one with a target at the nearest extremity and the other at the furthest; one with a target in the middle and the other in front; and one with a target in the middle and the other in the back.

Each participant then performed $2 \times 16 \times 4 = 128$ tasks, 32 of which were for training. The remaining 96 tasks were included in the analysis. Participants were given unlimited time to complete each task. Between each block, a screen indicated the progress of the experiment. Participants were then given the opportunity to take a break for as long as needed. Each session lasted approximately one hour, including instructions and post-experiment discussion.

5.4 Participants

We recruited 16 participants, with seven self-identifying as women and nine as men. Their ages ranged from 23 to 68 ($\mu = 37, \sigma = 16.2$, median = 29.5). Three participants were retired from various professions. The remaining participants were recruited within the laboratory and came from diverse fields. Among them, three were lecturers or researchers, and the other ten were PhD students or postdoctoral researchers. All participants had normal or corrected-to-normal vision, and none were color-blind. Seven participants reported being familiar with 3D virtual environments. Participation was voluntary and unpaid.

5.5 Measures

The experiment software recorded both the presented tasks and all participant manipulations: those involving the sliders used to report perceived depths, as well as interactions with the visualizations during the interactive phase. Additionally, we measured TIME, which corresponds to the duration between the moment the visualization was displayed and the moment when the participant pressed the space bar to validate their response. We collected these manipulations in log files for later analysis.

In the following, we characterize the relationship between the given instruction and the participant's perception using four metrics:

- *Order errors*, i.e. if the depth order of the two targets is reversed compared to the actual order;
- *Linearity of perception*, i.e. the relationship between perceived and actual depth of a target along the depth axis;
- *Judgement accuracy*, i.e. the difference between the perceived relative distance between the targets and the actual relative distance;
- *Judgment precision*, i.e. the spread of perceived depth judgments for a given stimulus.

Lastly, we analyze the effects of INTERACTION and DENSITY on these measures. INTERACTION indicates whether the visualization allowed user interaction to manipulate the angle of projection, and results below show that interaction is key to enabling users to make accurate judgments. DENSITY refers to the three levels of mark density used in the bubble charts. Since our investigations show that DENSITY has a negligible impact on the results, we do not consider it in the following section.

5.6 Results

5.6.1 Order Errors. To characterize D-MO as a visual channel, we first investigated whether it belongs to the group of channels capable of conveying a notion of order. According to Munzner, this quality distinguishes the *magnitude channels* (which allow the transmission of attributes with an inherent order) from the *identity channels* (which are suited for transmitting nominal attributes) [34]. To assess this, we counted the order errors, i.e. the number of tasks where the depths reported by the participants for the red and blue bubbles were in the reverse order of the actual one.

The mean error rate is 10.8 % (83/768) without interaction and 1.3 % (10/768) with interaction. Figure 7 shows how the errors are distributed according to the offsets between the targets. The smallest offset accounts for the highest number of errors, with a rate of

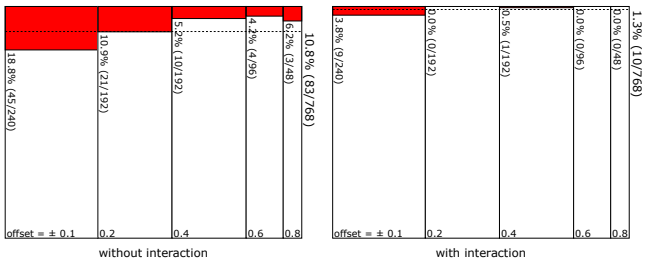

Figure 7: Rates and distributions of order errors, left: without interaction; right: with interaction. The rectangles represent the offsets between the targets. The widths of the rectangles are proportional to the number of tasks. Errors are represented by the red rectangles; the dotted line indicates the mean error rate.

18.8 % without interaction and 3.8 % with interaction. This error rate decreases as the offset between the targets increases.

We can compare this result with the errors reported by Heer and Bostock in their replication of the Cleveland and McGill experiments [13, 23]. They reported an error rate of 0.4 %, but the offsets they used in the experiments range from 0.18 to 0.83. This is comparable to our findings in the interaction condition for offsets greater than or equal to ±0.2: only one error was made among the 528 tasks, that is about 0.19 % errors, despite small offsets being overrepresented compared to large offsets.

Analysis at the participant level shows that errors are unequally distributed among them. In the condition without interaction, half of the errors were made by the three participants who committed the most errors. This suggests that the perception of depth conveyed by D-MO in animation mode does not work uniformly well for all individuals, as is the case with regular depth perception, i.e. stereoblindness [44]. However, this individual disparity disappears with interaction. This first result indicates that D-MO is well suited to convey information about attributes with an inherent order, provided interaction is available to users.

5.6.2 Linearity of Perception. To continue the characterization of D-MO, we investigated the relationship between perceived depth (denoted by p for perception) and actual depth (denoted by s for stimulus). An initial data exploration reveals a clear correlation between the two measures. The linearity of the relationship between p and s is one of Munzner's criteria for ranking visual channels, referred to as *accuracy* [34].

To assess this linearity, we performed a linear regression. Without interaction, the regression yields $p = 0.045{+}0.73{\times}s$, ($R^2 = 0.68$); whereas with interaction, it yields $p = 0.055 + 0.88 \times s$ ($R^2 = 0.80$). Once again, we observe individual disparities: participants with the most order errors have a substantially lower coefficient of determination (around $R^2 \approx 0.35$) compared to others (ranging from 0.7 to 0.95), both with and without interaction.

To fine-tune the analysis, we modeled the relationship between p and s using Stevens' law [52] since its exponent, α, is used as a criterion for ranking visual channels [34]. We used the direct judgment experimental formulation, in which two stimuli are presented and the participant is asked to estimate their ratio. In our case, for

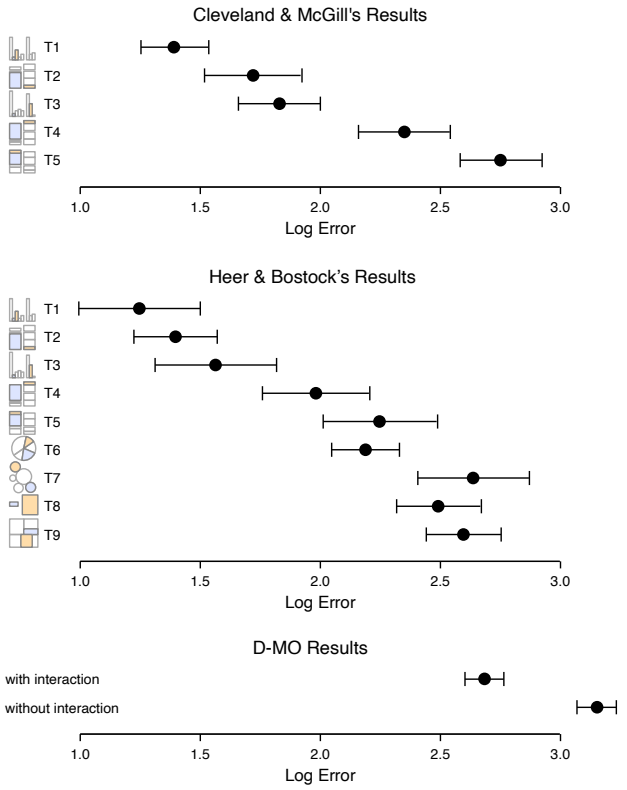

Figure 8: Proportional judgment (adapted from Heer & Bostock [23]), top: Cleveland & McGill's lab study [13]; middle: Heer & Bostock's studies [23] ; bottom: D-MO study. Error bars indicate 95% confidence intervals.

each target, we compared the perceived depth ratio p_r/p_b to the stimulus ratios s_r/s_b (where r and b correspond to the red and blue targets, respectively). Stevens' law is then expressed as:

$$\frac{p_r}{p_b} = \left(\frac{s_r}{s_b}\right)^{\alpha} \iff \log\frac{p_r}{p_b} = \alpha \times \log\frac{s_r}{s_b}.$$

A linear regression without intercept on the logarithms of the ratios allows us to estimate the exponent α in Stevens' law. These logarithms are defined only for positive values, which requires that both s_r and s_b have the same sign —that is, the two targets must lie on the same side of the origin. It is the case in half of the tasks (768 out of 1536), specifically those centered around $\{-0.8;\ -0.4; 0.4; 0.8\}$ with offsets of $\{\pm 0.1;\ \pm 0.2\}$. Additionally, p_r and p_b must also be of the same sign, which excludes other tasks, in particular those for which order errors occurred. The linear regression yields $\alpha = 1.04$ without interaction and $\alpha = 0.86$ with interaction. D-MO therefore performs slightly below the length (size in 1D) channel, for which $\alpha = 1$, but exhibits a perception closer to linearity than any other visual channel, such as area (size in 2D, $\alpha = 0.7$) or depth with perspective ($\alpha = 0.67$) [51].

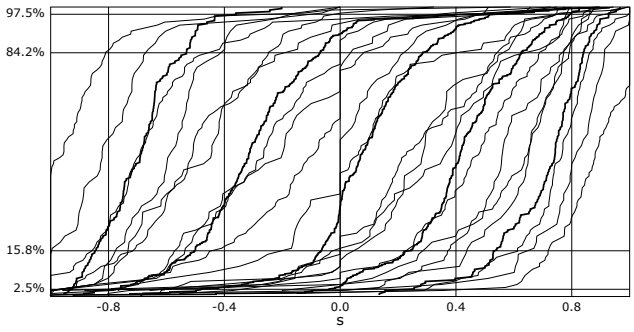

Figure 9: Cumulative distributions of p for each s, bold lines represent centers of the red/blue pairs while light lines represent each 21 individual levels of depth.

5.6.3 Accuracy. To quantify accuracy, Cleveland and McGill [13] introduced a measure called log *error*, based on the difference between the perceived ratio and the stimulus ratio. These ratios must be expressed as percentages to enable comparison with Cleveland and McGill's original study, as well as the replication by Heer and Bostock [23]. In our case, the log *error* is then defined as:

$$\log error = \log_2\left(|s_o - p_o| \times 100 + 1/8\right)$$

where s_o and p_o are the differences between the red and blue targets in the stimulus and the percept, respectively. We obtain $\log error = 3.15 \pm 0.08$ without interaction and 2.68 ± 0.08 with interaction ($\pm ci$ denotes 95 % confidence interval). Those values place D-MO with interaction among the length-based judgments, as shown on Figure 8, adapted from Figure 16 in Cleveland and McGill [13] and Figure 4 in Heer and Bostock [23].

Cleveland and McGill also quantified perceptual bias by analyzing the *large error rate*, defined as the proportion of trials with a log *error* greater than 4, which corresponds to a percent error exceeding 15.875 points. We obtain a large error rate of 33.2 % without interaction and 19.8 % with interaction. These results place D-MO between the position-based judgments (with at most 10 % of large errors) and length-based judgments (with at least 30 %, as shown in Figure 17 in Cleveland and McGill [13]).

5.6.4 Precision. Finally, we investigated D-MO's judgment precision, defined as the variability (or scattering) of percepts for a given stimulus. Figure 9 presents cumulative distributions of perceptions of perceived depth p as a function of stimulus depth s. The bold lines indicate the centers of the perception distributions, i.e. the mean of each red-blue target pair. There are five such centers, corresponding to the five possible central stimulus depths: $\{-0.8, -0.4, 0, 0.4, 0.8\}$. These depths are marked by vertical bars. The remaining light lines show the cumulative distributions for the red and blue targets considered separately. There are 21 such curves, corresponding to the 21 depth planes on which the targets are distributed (see Section 5.2 for a description of the tasks). The percentages marked on the vertical axis correspond to 1 standard deviation (50 ± 34.1 %) and 1.96 standard deviations (50 ± 47.5 % of the normal distribution, i.e. the bounds of the 95 % interval).

We observe that the distributions corresponding to the five median depths are mostly separated at 1 standard deviation, but show significant overlap at 1.96 standard deviations, which advocates for at most five distinct levels when using D-MO to encode categories.

5.6.5 Time. We then examined the impact of various parameters on task duration, referred to as TIME. TIME is defined as the time interval between the moment the task appeared on the screen and the moment the participant validated their response by pressing the space bar. Since participants tended to spend more time on more difficult tasks, understanding which factors influenced task duration can help refine the corresponding parameters.

First, we considered the effect of density (i.e. the number of bubbles) in the bubble charts. In our analysis, we separated the two phases of the experiment: with and without interaction. We then identified and removed outliers: 0.4 % in the non-interaction phase (durations exceeding 200 s) and 0.4 % of the data in the interaction phase (durations exceeding 150 s). Since TIME was not normally distributed in either phase, we applied a log transformation. A one-way ANOVA (analysis of variance) was conducted after checking the necessary assumptions. The results show no statistically significant differences across the three density levels: for the non-interaction phase $F_{2,762} = 1.060$, $p = 0.347$, with a very low proportion of variance explained ($R^2 = 0.003$); for the interaction phase, $F_{2,762} = 0.011$, $p = 0.989$, $R^2 = 2.75 \times 10^{-5}$. In conclusion, the time participants spent on the tasks was not influenced by the number of bubbles in the charts, as confirmed by the results of the one-way ANOVA conducted on the log-transformed TIME variable.

Secondly, we examined whether there is a relationship between TIME and the size of the targets (more precisely, their area). For each phase, we removed 0.4 % of outliers based on TIME (durations exceeding 200 s in the non-interaction phase and 150 s in the interaction phase). It is worth noting that, in this context, observations are organized in pairs: one observation corresponds to the red target, and the other to the blue target. Each pair represents a single task, and therefore, the recorded duration is the same for both observations. To account for this statistical dependence, we analyzed the red and blue targets separately. Scatter plots for both phases reveal dense and widely dispersed distributions, with no apparent trend. Additionally, the linear regression lines are flat, and the coefficients of determination are close to zero, indicating that target size does not account for the variability in TIME. Furthermore, a Spearman correlation analysis was conducted to explore a potential monotonic relationship between target size and TIME. The results indicate no significant correlation in the non-interaction phase for either the blue targets $\rho = 0.004$, $p = 0.914$ or the red targets $\rho = -0.030$, $p = 0.406$. Similarly, in the interaction phase, Spearman's coefficient for the blue targets does not reach statistical significance $\rho = -0.064$, $p = 0.078$, and no correlation is observed for the red targets $\rho = -0.018$, $p = 0.623$. These findings confirm the absence of a significant relation between target size and task duration, supporting the conclusion that size did not influence TIME.

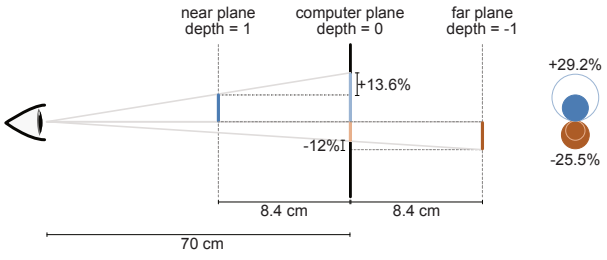

Figure 10: Size variation between projections with and without perspective. The nearest and farthest targets are positioned at ±8.4 cm from the central computer plane. The computer is positioned at 70 cm from the user's eyes. Under these conditions, the nearest target exhibits a 29.2 % increase in size when using perspective (13.6 % for its radius), and the farthest target exhibits a 25.5 % decrease in size (12 % for its radius).

6 Impact of D-MO on Size Channel

This section presents a second user study conducted to characterize the impact of D-MO on the size channel, which is typically the most affected one when introducing depth, especially when it relies on perspective. Since people are used to perspective in the real world, the perception of the sizes could be affected by D-MO: users may compensate for the distortion of the apparent sizes that would introduce perspective, despite D-MO relying on a parallel projection. In that case, users would overestimate the sizes of distant objects and underestimate the sizes of nearby objects. Figure 10 shows for the setup of the experiment the factors of distortion that would introduce a perspective projection: nearest (depth = 1) mark dimensions would increase by 13.6 % when projected on the screen plane, while farthest (depth = −1) mark dimensions would be reduced by 12 %. Since this factor affects the two dimensions of the plane, their 2D areas would thus be affected by a change ranging from −25.5 % to +29.2 %. If users compensate for this expected distortion that is not present with D-MO, their perception should underestimate the sizes of the marks close to them and overestimate the size of the marks far from them.

6.1 Visualization and Data

In this second experiment, we used the same type of bubble chart with the same attributes encoded by the same visual channels: the two standard components of position (x and y), depth based on D-MO, mark size, and color. We also used the same process described in Section 5.1 to generate random datasets for each task described below.

Following the same procedure as in the first experiment, the visualization was displayed within a square centered on the screen. Figure 11 shows the screen as presented during a task in this second experiment. The experiment was conducted in the same conditions as the first one.

6.2 Task

We employed a magnitude production method [48, 57] in which participants were asked to estimate the sizes of two targets selected from among the marks and highlighted in red or blue for each task. To record their judgments, two identical sets of adjustable bubbles were visible on the sides of the screen (Figure 11). Each set contained

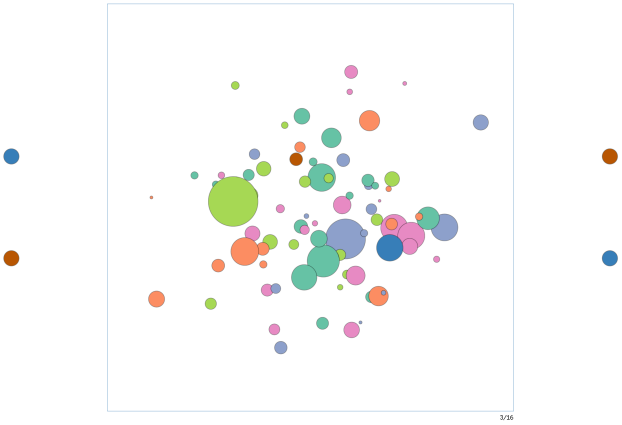

Figure 11: Example of a task in the second experiment. Participants must estimate the size of the red and blue bubbles by resizing the bubbles placed on both sides of the visualization to match their size.

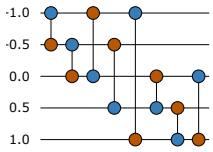

Figure 12: Depth configurations for size experiment.

one bubble corresponding to the blue and one to the red bubble. Participants were free to manipulate either set. Using the mouse wheel, they were asked to adjust the sizes of the two adjustable bubbles to match the perceived sizes of the red and blue bubbles, as displayed on the screen. Each tick of the mouse wheel increased or decreased the size of the bubble by a factor of $\sqrt[20]{2} \approx 103.5\,\%$, so that 10 ticks doubled the area of the mark. The adjustable bubbles could be adjusted as many times as needed, and in any order. To validate their response and proceed to the next task, participants pressed the space bar on the keyboard.

The depths of the two targets were set to five values within the interval $[-1, 1]$: $\{-1, -0.5, 0, 0.5, 1\}$. The eight depth combinations we used are shown in Figure 12.

A single density of 81 marks was chosen since the first experiment showed that density had no impact on D-MO. As in the first experiment, we ensured that the red and blue bubbles were easy to find by keeping them central and minimally occluded. However, rather than fixing a minimum size, we controlled their sizes. Four different sizes following a geometric progression were used: $\left\{2^{-3/4} \approx 0.59, 2^{-1/4} \approx 0.84, 2^{1/4} \approx 1.19, 2^{3/4} \approx 1.68\right\}$ times the average size of the bubbles present in the task. The size of the biggest target was thus eight times the size of the smallest one, and the range covered about 2/3 of the distribution of the sizes present in the tasks. We used the six possible pairs that could be built by picking two of those four sizes. These sizes covered a representative range of values shown on screen without excessively increasing the number of tasks in the experiment.

6.3 Procedure

The experiment followed the same procedure as the first experiment. First, participants were asked to fill in a consent form. Then, they were invited to sit in front of the computer running the experiment software.

The target configurations covered the eight combinations of depth times the six pairs of sizes, leading to a total of 48 tasks. As in the first experiment, these tasks were split into three blocks of 16 tasks each, and an additional training block of 16 tasks was first presented. The training block began with the four tasks involving the smallest and largest sizes and depth. In this second experiment, participants were always allowed to use the interactions described in Section 4.2, since we had found that D-MO performed best with interaction.

Each participant performed $16 \times 4 = 64$ tasks, 16 of which were for training purposes and excluded from the analysis. We thus had $48 \times 2 = 96$ data points per participant, and $96 \times 16 = 1536$ points overall.

As in the first experiment, participants were given unlimited time to complete each task. Between each block, a transition screen indicated the progress of the experiment. Participants were then given the opportunity to take a break for as long as needed. Each session lasted approximately 45 minutes, including instructions and post-experiment discussion.

6.4 Participants

We recruited 16 participants who had not taken part in the first experiment, with seven self-identifying as women and nine as men. Their ages ranged from 17 to 41 ($\mu = 25.2, \sigma = 7.3$, median = 23.5). The participants were recruited within the laboratory and came from diverse fields. Three of them were research engineers, and one was a university lecturer. The remaining participants were interns, PhD students, or postdoctoral researchers. All participants had normal or corrected-to-normal vision, and none were color-blind. Seven participants reported being familiar with 3D virtual environments. Participation was voluntary and unpaid.

6.5 Measures and Results

As in the first experiment, the software logged all user interactions. For this experiment, we concentrated our analysis on the specific question of the impact of depth on the perceived sizes of the marks. The targets were characterized by their radius s and their depth z. For both targets of each task, we computed r, the ratio between their perceived radius p and their actual ones s. We thus had $48 \times 2 = 96$ data points per participant, and $96 \times 16 = 1536$ points overall.

6.5.1 Factors impacting Size Perception. An ANOVA modeling the effects on the perceived size p shows that two factors have significant contributions: the actual size s and the depth z. The impact of actual size is obvious and indeed very strong, which is why we examine the p/s ratio below. We designed the tasks so that each mark size appeared the same number of times at every depth level, ensuring that the effects of s and z were independent. The ANOVA confirms that there is no significant interaction between s and z.

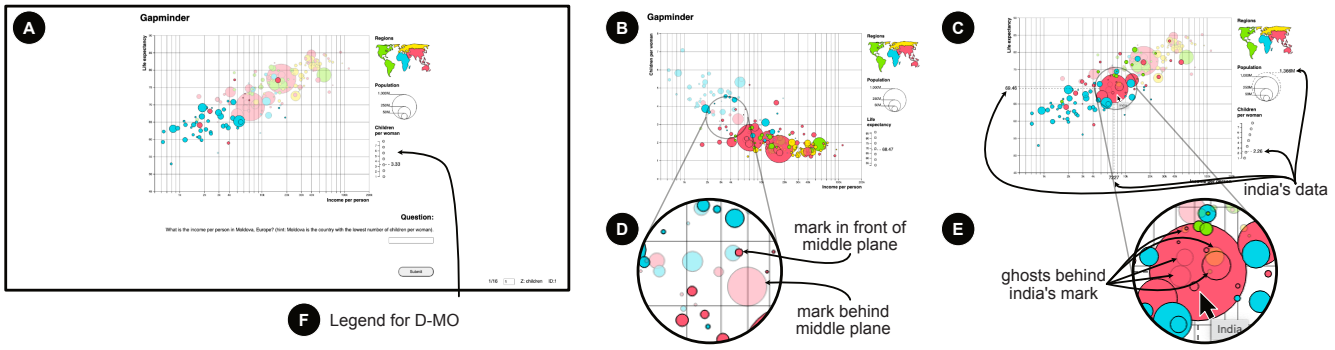

Figure 13: Bubble Chart with D-MO. A) Screenshot of the pre-study screen with the visualization and a question. B) Visualization when the D-MO and the Y-axis mappings are exchanged. C) Hovering over a mark displays its name under the cursor and its data on the corresponding legend. D) Marks behind the central plane of movement for D-MO appear less saturated as the plane is white and transparent. E) Marks behind another one appear as ghosts. F) Legend for D-MO. The central plane value is written on the right.

6.5.2 Effect of Depth on Size Perception. A linear fit of the 1536 data points provides the following model:

$$r = 1.0067152 - 0.0168142 \times z.$$

The RMSE is low (0.072), indicating that z explains only a small part of the variability in r. However, the t-tests for the slope show that it is significantly non-zero ($t = -6.31$, $p < 0.0001$). The negative slope is expected, as explained at the beginning of Section 6: users tend to overestimate the size of distant marks and underestimate the size of closer marks. Since z ranges from -1 to 1, r ranges from 0.989 to 1.024, which—when squared to account for the 2D size of the marks—corresponds to a maximal underestimation of −2 % and overestimation of +4.8 % of their areas.

We can compare these values to the −25.5 % to +29.2 % that would be expected if users fully corrected for a perspective that is not present. The distortion introduced by D-MO in the perception of size is far smaller than this worst-case scenario. We can also compare these values with the perceptual bias already present for 2D areas: observers tend to underestimate large areas. As noted earlier, this bias is characterized by Stevens' law with an exponent of 0.67 [51]. In our experiment, the area of the largest target was eight times the area of the smallest. Their perception would then be in the ratio of $8^{0.67} \approx 4$. This implies that a factor of two is missing in the perception. Such a factor corresponds to an underestimation of 50 % or an overestimation of 100 %. Despite this substantial distortion, using 2D size is pretty common in visualizations.

We also performed a linear fit of the data per participant to check if there was significant variability among people. The intercepts range from 0.972 to 1.052, with a median of 1.003. The slopes are all negative and range from −0.036 to −0.004, with a median of −0.012. For half of the participants, the t-test does not confirm that the slope is significantly non-zero.

We can conclude that people tend to correct slightly for a perspective deformation that is not present, but that this correction is relatively small compared to other biases affecting the perception of size.

7 D-MO in Use as a Visual Channel

This section presents a pre-study where D-MO is implemented in an information visualization. The aim of this pre-study is to test whether users can accurately read and interpret a visualization that uses D-MO, rather than assessing the effectiveness of D-MO compared to other visual channels. Additionally, this pre-study proposes a first legend design for D-MO, as well as additional interaction to facilitate the use of D-MO.

7.1 Visualization and Data

As in our previous studies, we used a bubble chart. We used a dataset made publicly available by Gapminder [17], an independent non-governmental organization dedicated to fighting global misconceptions [16]. Furthermore, the design of our bubble chart drew inspiration from the visualization tool employed by Gapminder [15]. This visualization is illustrated in Figure 13A and the accompanying video. Each mark (or *bubble*) represents a country worldwide. For all visualizations in this pre-study, color was used to encode the respective world regions, the X-axis encoded the GDP per capita, and area (2D size) encoded population. In the training visualization, the Y-axis encoded a happiness score, while D-MO encoded forest coverage. For the experiment visualization, the Y-axis and D-MO encoded, respectively, the life expectancy and the number of children per woman, or the opposite, depending on the evenness of the user's running rank (see Figure 13A and B). This enabled us to compare the performance of D-MO with that of the Y-axis.

In addition to the design, we also implemented similar interaction mechanisms as those used by the original Gapminder's visualization tool [15]. When hovering over a data point, the corresponding value for each axis of the bubble chart was displayed in the legend, as illustrated in Figure 13C and the accompanying video. Additionally, the name of the country was displayed in a pop-up next to the bubble. Upon clicking on a region within the color legend, the visibility of the bubbles belonging to that region was toggled.

Finally, for D-MO, we retained the same interactions as in our previous experiments. Hovering over a bubble made its depth the

reference for the movement of other points. Dragging a bubble enabled direct control of the projection axis, and in turn, of the bubble movements. In addition, we introduced four novel features for interacting with D-MO. The first feature was a materialization of the reference plane by adding a grid stemming from the X- and Y-axis ticks. The grid was completed by a white and slightly transparent rectangle, creating a difference in saturation for bubbles in front of the plane, which retained their normal color, and those behind the plane, which appeared less saturated (see Figure 13D). The second feature was the display of ghosts for hidden marks. Because occlusion is a strong cue for D-MO perception, marks were drawn in depth order instead of reverse size order, as is normally done in bubble charts. To enable users to see and interact with these hidden marks, a ghost was displayed on top of the other data points (see Figure 13E). The interaction with a ghost was equivalent to the interaction with a mark. These ghosts were inspired by the slightly transparent marks used in Gapminder's bubble chart [15] to avoid occlusion. The third feature was a legend for D-MO. The legend mapped the depth range onto a vertical axis, with dots having their motion controlled by the corresponding depths along the axis (see Figure 13F). Hovering over the legend set the depth reference to the position of the cursor within the legend. The fourth feature disabled D-MO's animation when the cursor was positioned outside the visualization or its legend. This prevented the movements of marks from interfering with the reading of values on the X-axis.

The interactive visualization is available in the supplementary material.

7.2 Task

Participants were asked to answer questions about the displayed visualization. The questions were displayed at the bottom of the screen, with only one question visible at a time. Participants interacted with the visualization using a mouse, and answered the questions in the dedicated form located next to the question. To validate their answer, participants were required to click on a submit button displayed on the screen.

Two sets of questions were designed: one for the training phase and one for the experimental phase. The training phase set consisted of four questions, each focusing on reading either D-MO or Y-axis values (see Table 1). The experimental phase set comprised sixteen questions, covering all visual channels of the visualization, to ascertain whether D-MO influenced users' ability to correctly answer questions about the other visual channels (see Table 2).

By putting our task abstraction in parallel with Munzner's action levels [34], the main mid-level actions exploited here were *browse* and *explore*. Only a few questions required *locate*. At Munzner's lower level, our questions corresponded to the three scopes: *identify*, *compare*, and *summarize*. According to Amar et al.'s taxonomy [1], the majority of our questions aligned with the tasks *Retrieve value*, *Filter*, and *Find extremum*. Finally, in our preliminary study, the main tasks given to the participants were *identification tasks*, more precisely, *absolute-value identification* (Questions 2-8, 10, 13-16) and *relative-value identification* (Questions 1-4, 9, 10, 12-14), as introduced by Szafir et al. [53], since participants had to locate countries with given attributes or find countries relative to the distribution. The second most common tasks were *segmentation tasks*, more specifically, *segmentation by features* (Questions 2-4, 8, 9, 14-16), since participants, when given a specific region, had to consider only countries with the corresponding color. A few questions targeted *structure estimation tasks*, specifically *similarity detection* (Questions 9, 11, 14), when determining similarity across regions; and, *trend detection* (Question 11). Finally, fewer questions concerned *summary tasks*, especially *numerosity* (Questions 6, 15), as participants had to count the number of countries with matching attributes; and *mean and variance estimation* (Question 9). As we can see, we did not include questions requiring *outlier identification*, nor did we ask about the distribution of the bubble chart (e.g., *distribution statistics* or *distribution shape*).

Table 1: Set of training-phase questions and the visual channels required to answer them.

Question	Visual channel(s) required to answer
In Asia, which country has the highest happiness score?	Y
Which country has a happiness score of 69.44?	Y
In Europe, which country has the highest forest coverage?	D-MO
Which country has a forest coverage of 23.93?	D-MO

7.3 Procedure

First, participants were required to complete a consent form that provided a comprehensive overview of the experimental objectives. Subsequently, they were invited to proceed to a computer equipped with the experimental software.

The experimenter first demonstrated the potential interaction with the visualization. Then, participants started the training phase. During this phase, participants were allowed to ask any questions regarding the interaction with the visualization. Upon completion of the training phase, participants moved on to the experimental phase. At the beginning of the experimental phase, the experimenter reiterated the primary objective of the experiment, which was to accurately interpret the visualization without excessive time allocation. During the experimental phase, verbal comments were recorded by the experimenter.

For both phases, the order of the questions presented was randomized. Each session lasted approximately 45 minutes, encompassing instructions and a post-experiment discussion. After the experiment, participants were given the possibility to provide any comment they had on the visualization.

7.4 Participants

We recruited six participants familiar with Information Visualization: three use visualizations on a daily basis, one on a weekly basis, and two on a monthly basis. Their ages ranged from 24 to 41 ($\mu = 30.5$, $\sigma = 6.4$, median $= 28$). The participants were interns or PhD students in computer science. They all had normal or corrected-to-normal vision, and none were color-blind. Participation was voluntary and unpaid.

Table 2: Set of experiment-phase questions and the visual channels required to answer them.

Id	Question	Visual channel(s) required to answer	
		D-MO mapped to children per woman	D-MO mapped to life expectancy
1	Across all four world regions, which country ranks third in terms of population?	size	
2	In Africa, which country ranks first in terms of population?	size, color	
3	Which country in the Americas has the lowest income per person?	X, color	
4	In Africa, which country has the highest life expectancy?	Y, color	D-MO, color
5	Which country has a life expectancy of 61.62 years?	Y	D-MO
6	How many countries have a life expectancy lower than 60 years?	Y	D-MO
7	Is there at least one country with a life expectancy of exactly 75.00 years?	Y	D-MO
8	Which country(ies) in the Americas has(have) an income per person ranging between 9 k€ and 10 k€?	X, color	
9	Which region shows the highest variation in the number of children per woman?	D-MO, color	Y, color
10	What is the income per person in Portugal, Europe? (hint: Portugal is the country with the lowest number of children per woman with Taiwan).	X, D-MO	X, Y
11	Do you observe a trend between the life expectancy and the number of children per woman across all four world regions?	Y, D-MO	D-MO, Y
12	Is the country with the highest life expectancy located in the same region as the country with the highest income per person?	Y, X, color	D-MO, X, color
13	Compare the number of children per woman in Brazil and China. (hint: these countries rank second and first in terms of population in the Americas and Asia, respectively).	D-MO, size, color	Y, size, color
14	Which region shows the largest gap between the country with the lowest and the country with the highest life expectancy?	Y, color	D-MO, color
15	How many countries in Africa have more children per woman than Nigeria?	D-MO, color	Y, color
16	Which country(ies) in Africa has(have) a lower income per person than Niger? (hint: Niger has the highest number of children per woman).	X, D-MO, color	X, Y, color

7.5 Measures

The experimental software was configured to record both the presented questions and all participant interactions with the visualizations. We measured CORRECTNESS, which characterizes whether participants answered the question correctly. Additionally, we measured TIME, which corresponds to the duration between the display of the question and the validation of the answer with a click on the button by the participant. We also analyzed the effects of VISUAL CHANNEL on these measures.

7.6 Results

7.6.1 Correctness. We recorded a total of 96 questions (6 participants × 16 questions), excluding the training questions. 93 % (89/96) of the questions were correctly answered. Out of the 7 errors, three were identified as question-reading errors. Two of these three errors involved the participant correctly answering the number of countries instead of their names. The third error involved the participant correctly answering a question about the X-axis instead of the Y-axis. Table 3 shows the errors per question.

Of the remaining errors, two were related to reading Y-axis values (Question 14), and the other two concerned reading the D-MO values (Question 13). We believe that the errors for the Y-axis values may be due to the occlusion of one mark required to answer the question. For the D-MO values, the question required identifying the minimum and maximum life expectancy values for each region and then computing the difference. Two regions had very similar differences: Asia (26.2 years) and Africa (25.82 years). From our observations during the experiment, users who answered incorrectly did not perform the calculation, but instead attempted to visually deduce the answer. These errors might not be dependent on D-MO, but rather on the user's answering strategy.

Table 3: Count of errors per question, data mapped to D-MO, and whether the question required reading D-MO to answer. Only questions for which at least one user answered incorrectly are included. Question errors are those caused by incorrectly reading the question. Visual channel errors are those caused by incorrectly interpreting the visual channel.

Id	D-MO mapping	Requires D-MO	Total errors	Question errors	Visual channel errors
8	Children per woman	False	1	1	0
8	Life expectancy	False	0	0	0
14	Children per woman	False	0	0	0
14	Life expectancy	True	2	0	2
15	Children per woman	True	0	0	0
15	Life expectancy	False	3	1	2
16	Children per woman	True	1	1	0
16	Life expectancy	False	0	0	0

7.6.2 Time. Due to the limited number of data points, we were unable to perform any statistical tests. Figure 14 shows the response time per question, data encoded by D-MO, and whether the question required reading D-MO to answer. Participants took 100.7 s (σ = 53.9) to answer questions requiring reading D-MO, compared to 71.2 s (σ = 46.4) when requiring reading the Y-axis. However, the ratio between both visual channels varies greatly, with some questions answered twice as fast when using D-MO while others were twice as fast when using the Y-axis.

7.6.3 Qualitative comments. Based on users' comments, it seems that using D-MO was more challenging than using the X- or Y-axis, but comparable to size. Specifically, P1 and P4 mentioned that it was difficult to find a specific value using D-MO, while P2, P3, P4, and P5 noted that it was challenging to compare sizes. P4 and P5 emphasized that the difficulty for both primarily stemmed from the density of the plot. P4 mentioned that hovering over small targets was difficult, especially when they were moving. P3 used the drag interaction in D-MO to compare sizes by trying to bring data points closer together.

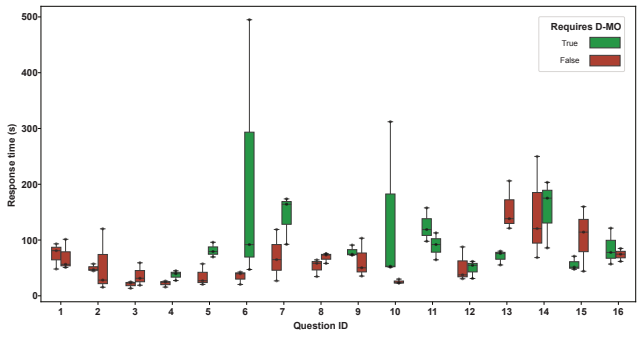

Figure 14: Box plot of the response time per question, data mapped to D-MO, and whether the question required reading D-MO to answer. For each question ID, the left box represents the mapping of children per woman to D-MO, while the right box represents the mapping of D-MO to life expectancy.

8 Discussion and Future Work

8.1 Characteristics of D-MO

Munzner proposes a set of characteristics for evaluating the *expressiveness* and *effectiveness* of visual channels [34]. In this section, we present the characteristics of D-MO according to this list. We also discuss how these characteristics align with Bertin's visual channel properties [2].

8.1.1 Expressiveness. Munzner proposes the following definition for *expressiveness*: a visualization must represent all attribute information from the dataset without introducing any extraneous information [34]. According to this principle, a visualization should convey a perception of order only when the underlying data is ordered, and avoid implying order when the data has no meaningful order [34]. *Expressiveness* corresponds to Bertin's notion of *order* [2].

The analysis of order errors in Section 5.6.1 showed that D-MO effectively conveys information about the order of targets with a low error rate when interaction is enabled. This supports D-MO's *expressiveness* and places it among the *magnitude* visual channels, which are suitable for encoding attributes with an inherent order.

8.1.2 Effectiveness – Separability. According to Munzner, one of the components of effectiveness is separability, i.e. how the level of a visual channel influences the perception of the other visual channels used, and vice versa. This echoes Bertin's notion of *association*. In a pilot study, we identified that the perception of D-MO improves when the color and size channels are also used. This is generally not the case for visual channels, whose perception tends to decrease as the number of channels used increases. We hypothesize that these two channels are beneficial because they reinforce occlusion cues. Indeed, size increases the probability of occlusion occurring and the diversity of their contours, while color facilitates the identification of partially covered marks.

Conversely, our second study shows that D-MO slightly influences the perception of size. As shown in Section 6.5.2, the under- or overestimation of the size is just by a few percent, whereas the bias of 2D size perception on its own or the distortion introduced by perspective can reach more than 25 %. We, thus, consider that this drawback can be accepted since the benefit of D-MO is to extend the very short list of visual channels that can provide order and quantity judgments.

We also hypothesized that D-MO might affect the interpretation of the X-axis due to the movements along that axis. However, the animation pauses at the central position, and interaction allows users to freeze a specific target by hovering over it or the entire visualization by moving the cursor outside the visualization. We designed D-MO animation and interaction to mitigate its impact on the horizontal position. Our pre-study testing D-MO in a visualization showed that users could accurately interpret the X-axis when D-MO was used. In fact, no users reported difficulties reading the X-axis. We hypothesize that adding the interaction to disable D-MO, i.e. when the cursor is not above the visualization, helped participants compare X-axis values. However, D-MO could influence reading an X-axis value for small marks. In fact, reading an X-axis value was done by hovering over a mark to display its X value on the axis, and one participant reported difficulties hovering over smaller targets.

The exact measurement of the influence of D-MO on other visual channels is beyond the scope of this paper and left to future work.

8.1.3 Effectiveness – popout. Visual popout refers to the ability to immediately identify an object among other visual elements in a pre-attentive manner [34].

In a configuration where the majority of marks are located near the central depth plane, an outlier placed in an extreme position (e.g., at the front) will be encoded by D-MO with high speed and amplitude of movement, and will therefore be immediately identifiable, as its depth will clearly stand out from the others. Two of the underlying principles are the facts that movement is a particularly salient visual cue [34] and popout stimulus [21]. Gutwin et al. [21] found that motion showed high popout accuracy, even at low intensities and at large visual angles.

Furthermore, as shown in Section 5.6.5, under conditions of different densities and mark sizes, their influence on DURATION is not noticeable. This seems to indicate the pre-attentive nature of D-MO. However, as with separability, we leave the study of this characteristic for future work.

8.1.4 Effectiveness – grouping. Grouping refers to the perception of certain elements as belonging to the same group due to common visual characteristics. This aligns with Bertin's concept of *selection*. In the context of D-MO, the perception of these groups is partly justified by the literature: objects moving at the same speed are known to be perceived as belonging to the same group [31].

We leave the characterization of this component for future work. However, this notion is strongly linked to the notion of discriminability presented below.

8.1.5 Effectiveness – discriminability. Discriminability refers to the ability of a visual channel to make differences between represented values perceptible. This aligns with Bertin's concept of *selection*, and specifically with the number of classes that can be formed using the visual channel.

Our first controlled experiment shows that five levels of depth are mostly separated (see Section 5.6.4). Moreover, the significant

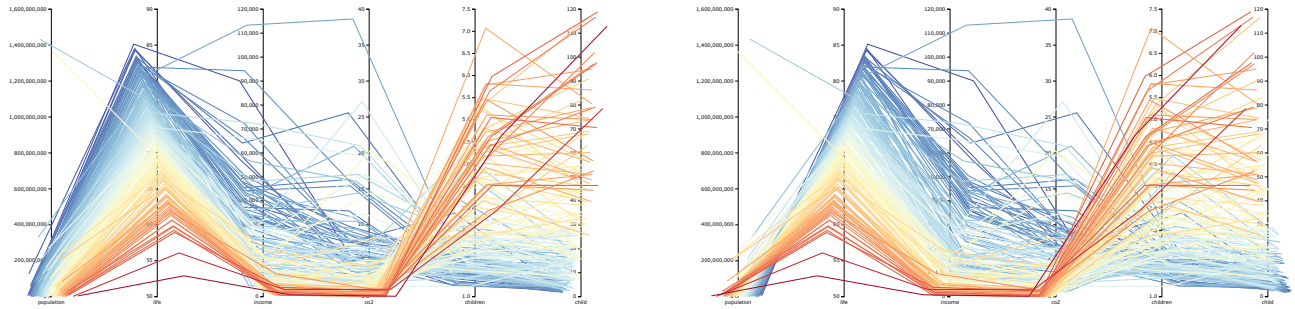

Figure 15: Parallel Coordinates with D-MO, left and right projections. The vertical axis encodes different indicators (population, life expectancy, income, etc.) D-MO encodes the indicator of the hovered axis, here the second from the left.

decrease in error rate as the offset between targets exceeded ±0.1 supports this range. This demonstrates that D-MO provides limited discriminability; however, five levels remain useful: Bertin gives the same upper bound of five levels for channels capable of conveying at least *order*, i.e. size, value and texture. To effectively encode groups by distinct levels, D-MO should be restricted to a maximum of five depths.

8.1.6 Effectiveness – accuracy. Finally, accuracy refers to the user's ability to estimate the value of the data represented. In other words, it is the difference between the value perceived by the user and the value of the stimulus. This concept aligns with Bertin's idea of *quantity*. Poor fidelity can result in an overestimation or underestimation of the actual value of the data. Fidelity can be characterized by Stevens' psychophysical law [52], which states that the perceived value (S) increases as a power function of the intensity of the objective value (I): $S = k \times I^{\alpha}$ (where k is a constant and α is an exponent varying between 0.5 and 3.5 depending on the channel).

As indicated in Section 5.6.2, the exponent of Stevens' law without interaction is $\alpha = 1.04$, and $\alpha = 0.86$ with interaction. Since both values are very close to 1, we can conclude that the depth perception induced by D-MO follows its variations almost linearly. This linearity applies only to the range of depth tested in the experiments, which is of the same order as the range available for the X- and Y-axes. This linearity is nonetheless noteworthy and makes D-MO better than 2D size or actual depth regarding this criterion, confirming its perceptual effectiveness.

8.2 D-MO for Information Visualization

From our third experiment (pre-study), it appears that D-MO is usable in visualizations. As reported in Section 7.6.1, users were able to accurately answer 93 % of the questions on a bubble chart. However, there seems to be a potential impact of D-MO due to the drawing order required by the occlusion, which was responsible for two out of four errors. Finding a better design for the ghosts and the interaction with marks occluded could probably improve D-MO.

In addition, further consideration should be given to the design of the interaction and the legend of D-MO in a visualization. P3 and P6 appreciated the possibility to stop the movement, but would have preferred to stop it with a button to enable further exploration of the visualization with the cursor without restarting D-MO. For the legend, P1 and P6 would have preferred to flip the direction of the D-MO legend, so that higher values are further in depth. Additionally, P6 would have liked to be able to lock the central plane of movement from the legend. Despite these possible explorations for future work, the current legend is a first viable step. Users expressed how useful D-MO is, particularly for identifying trends (P2 and P6).

8.3 Augmenting Visualizations with D-MO

We also implemented alternative visualizations to bubble charts using D-MO, as shown in the video figure. As expected from the perception literature [5], D-MO is difficult or impossible to perceive in visualizations without occlusion or overlaps of many marks, such as bar charts or pie charts. However, we present here some examples where it works well.

D-MO can complement parallel coordinates [25] by redundantly encoding the axis hovered over by the mouse. Then, the correlations between this axis and other distant axes, which are known to be hard to perceive in the original visualization, are easier to see. On Figure 15, we show parallel coordinates augmented with D-MO applied to the Gapminder dataset [17].

D-MO can also be used in conjunction with terrain visualization to show digital elevation model (DEM). It then provides a perception similar to the relief shearing introduced by Willett et al. [65], but with a different interaction.

With the bubble chart used in the experiments, we present three very different examples: one with point marks (bubbles), another with line marks (parallel coordinates), and the last one with area marks (terrain). This shows the versatility of D-MO and its potential to expand the set of visual channels available for a variety of visualizations.

8.4 Limitations

D-MO relies on two depth cues: occlusion and motion.

For occlusion to function effectively, marks need to be drawn in depth order rather than in the size order typically used in bubble charts to avoid occlusion. In our first two studies, we avoided complete occlusion of the targets to ensure that the time required to find a target did not vary from one task to another. In our pre-study, we implemented ghosts for hidden marks, which are transparent versions of the hidden mark displayed above the hiding mark. However, a complete study is needed to quantify how these ghosts

perform for hidden marks compared to when they are not hidden. Furthermore, we only used filled circles as marks. However, D-MO could be less efficient with other mark shapes or hollow shapes, as there may be less occlusion.

Motion can be used to visualize categorical data [34] or the evolution of data over time, such as in Gapminder's bubble chart [15]. In our pre-study, we used Gapminder's data [17] from a specific date to avoid conflicts between the animation over time and the animation of D-MO. It is possible to envision using both in parallel by allowing users to switch between visualizing over time or for a specific date. As with other visual channels, due to the scarcity of visual channels, designers need to carefully select which visual variable to map to which data attributes.

In this paper, we focused on studying the visual encoding design level [34] and only performed a pre-study for the data/task abstraction design level [34]. A study with more participants and covering more visualization tasks is required to understand how D-MO performs in a visualization. Our pre-study proposed initial results and an initial design to integrate D-MO into an information visualization. Based on participant comments, users only encountered difficulties when locating a specific value. They reported that this difficulty was similar to finding a specific size. This issue could be solved by highlighting the points on the middle plane when interacting with D-MO's legend or adding filtering capabilities.

9 Conclusion

In this article, we present D-MO, a novel visual channel that employs a perspective-free approach to convey depth. D-MO is based on a fusion of visual cues, namely relative motion and occlusion, whose impact on depth perception has been substantiated through experimental psychological research. Relative movements are generated by an oblique projection that preserves the size of visual marks. Two complementary modes are available: an animation that continuously modulates the projection axis between four distinct viewpoints (left, center, right, and center); and an interaction that enables the user to directly manipulate the origin and the orientation of the projection axis.

D-MO does not necessitate any specialized devices or equipment, facilitating seamless use for visualization. Furthermore, the introduced visual channel is effective. In fact, our controlled experiments demonstrate that D-MO ranks among the most expressive visual channels, those capable of encoding quantitative attributes. Based on the various criteria employed for classifying these channels (i.e. ability to convey an order relationship, Stevens' exponent, Cleveland and McGill's log $error$, and Bertin's properties), D-MO falls within the category of position and size channels. While its discriminability is not high, it falls within the acceptable range of Bertin's other selective channels (e.g., for punctual marks: four levels for size, three for value). The significant proportion of errors attributed to a limited number of participants indicates that D-MO may not be universally effective, as is also the case with regular depth perception. Consequently, further investigation is required to determine whether specific parameters, such as animation frequency and amplitude, need customization to optimize perception for each individual. Moreover, as our second experiment shows, D-MO impacts size perception marginally.

D-MO can be incorporated into visualizations. Our pre-study of a bubble chart visualization augmented with D-MO and interactions similar to those in GapMinder [15], indicated that participants were able to answer questions, whether or not they required reading D-MO values. We proposed a preliminary design for the legend of D-MO and a preliminary mechanism to address the complete occlusion of some marks. Although further study and evaluation on these elements are necessary, the proposed design is already usable.

Finally, D-MO can be implemented in other types of visualizations, as shown in our discussion, and we believe that this paper will facilitate the conception and implementation of novel visualizations with a usable and efficient depth visual channel.

Acknowledgements

This work was partly supported by the French National Research Agency (ANR) project SOLAS (ANR-23-IACL-0006).

References

[1] Robert Amar, James Eagan, and John Stasko. 2005. Low-level components of analytic activity in information visualization. In *Proc. IEEE InfoVis 2005*. IEEE, 111–117.
[2] Jacques Bertin. 1967. *Sémiologie graphique.* Mouton/Gauthier-Villars.
[3] Anastasia Bezerianos and Petra Isenberg. 2012. Perception of Visual Variables on Tiled Wall-Sized Displays for Information Visualization Applications. *IEEE TVCG* 18, 12 (2012), 2516–2525. doi:10.1109/TVCG.2012.251
[4] Myron L. Braunstein. 1962. Depth perception in rotating dot patterns: effects of numerosity and perspective. *J. Exp. Psychol.* 64, 4 (1962), 415.
[5] Myron L. Braunstein. 1962. The perception of depth through motion. *Psychological Bulletin* 59, 5 (1962), 422.
[6] Cynthia Brewer. 2018. ColorBrewer: color advice for cartography. http://www.ColorBrewer.org
[7] Athena Buckthought, Ahmad Yoonessi, and Curtis L. Baker. 2017. Dynamic perspective cues enhance depth perception from motion parallax. *J. Vis.* 17, 1 (2017), 10.
[8] Kelly Caine. 2016. Local Standards for Sample Size at CHI. In *Proc. ACM CHI'16*. 981–992. doi:10.1145/2858036.2858498
[9] Stuart K. Card and Jock Mackinlay. 1997. The Structure of the Information Visualization Design Space. In *Proc. InfoVis'97*. 92–99.
[10] David L. Carpenter. 1979. Development of depth perception mediated by motion parallax in unidimensional projections of rotation in depth. *J. Exp. Child Psychol.* 28, 2 (1979), 280–299.
[11] Corrado Caudek and Dennis R. Proffitt. 1993. Depth perception in motion parallax and stereokinesis. *J. Exp. Psychol.: Human Perception and Performance* 19, 1 (1993), 32.
[12] Ed H. Chi and John T. Riedl. 1998. An Operator Interaction Framework for Visualization Systems. In *Proc. InfoVis'98*. 63–70.
[13] William S. Cleveland and Robert McGill. 1984. Graphical perception: Theory, experimentation, and application to the development of graphical methods. *J. Am. Stat. Assoc.* 79, 387 (1984), 531–554.
[14] Jocelyn Faubert. 2001. Motion parallax, stereoscopy, and the perception of depth: Practical and theoretical issues. In *Three-Dimensional Video and Display: Devices and Systems: A Critical Review*, Vol. 10298. SPIE, 168–191.
[15] Gapminder. Accessed on November 20, 2025. *Bubble chart visualisation tool.* www.gapminder.org/tools/#\protect\T1\textdollarchart-type=bubbles&url=v2
[16] Gapminder. Accessed on November 20, 2025. *Gapminder.* www.gapminder.org
[17] Gapminder. Accessed on November 20, 2025. *Gapminder's data, free to user CC-BY Gapminder.* https://www.gapminder.org/data/
[18] Wenjing Geng, Ran Ju, Xiangyang Xu, Tongwei Ren, and Gangshan Wu. 2015. Flat3D: browsing stereo images on a conventional screen. In *Proc. Int. Conf. on Multimedia Modeling*. Springer, 546–558.
[19] Eleanor J. Gibson, James J. Gibson, Olin W. Smith, and Howard Flock. 1959. Motion parallax as a determinant of perceived depth. *J. Exp. Psychol.* 58, 1 (1959), 40.
[20] Claire C. Gordon, Cynthia L. Blackwell, Bruce Bradtmiller, Joseph L. Parham, Patricia Barrientos, Stephen P. Paquette, Brian D. Corner, Jeremy M. Carson, Joseph C. Venezia, Belva M. Rockwell, et al. 2014. 2012 anthropometric survey of us army personnel: Methods and summary statistics. *Army Natick Soldier Research Development and Engineering Center MA, Tech. Report* (2014).

[21] Carl Gutwin, Andy Cockburn, and Ashley Coveney. 2017. Peripheral popout: The influence of visual angle and stimulus intensity on popout effects. In *Proc. ACM CHI'17*. 208–219.

[22] Douglas A. Hanes, Julia Keller, and Gin McCollum. 2008. Motion parallax contribution to perception of self-motion and depth. *Biological Cybernetics* 98 (2008), 273–293.

[23] Jeffrey Heer and Michael Bostock. 2010. Crowdsourcing Graphical Perception: Using Mechanical Turk to Assess Visualization Design. In *Proc. ACM CHI'10*. 203–212.

[24] Paul B. Hibbard, Alice E. Haines, and Rebecca L. Hornsey. 2017. Magnitude, precision, and realism of depth perception in stereoscopic vision. *Cognitive Research: Principles and Implications* 2 (2017), 1–11.

[25] Alfred Inselberg. 1985. The plane with parallel coordinates. *The Visual Computer* 1, 2 (1985), 69–91.

[26] Mary K. Kaiser and Dennis R. Proffitt. 1992. Using the Stereokinetic Effect to Convey Depth: Computationally Efficient Depth-From-Motion Displays. *Human Factors* 34, 5 (1992), 571–581.

[27] J. M. Loomis and D. W. Eby. 1989. Relative motion parallax and the perception of structure from motion. In *Proc. Workshop On Visual Motion*. IEEE Computer Society, 204–205.

[28] Scott MacKenzie. 2013. *Human-Computer Interaction: An Empirical Research Perspective*. Morgan Kaufmann, Chapter 5.

[29] Jock Mackinlay Mackinlay. 1986. Automating the Design of Graphical Presentations of Relational Information. *ACM Trans. Graph.* 5, 2 (1986), 110–141.

[30] John P. McIntire and Kristen K. Liggett. 2014. The (possible) utility of stereoscopic 3D displays for information visualization: The good, the bad, and the ugly. In *Proc. IEEE 3DVis'14*. 1–9. doi:10.1109/3DVis.2014.7160093

[31] Albert Michotte. 2017. *The perception of causality*. Routledge.

[32] Frederick R. Mosher. 2017. Wiggle 3D Displays of Weather Data. (2017).

[33] Tamara Munzner. 2009. A Nested Model for Visualization Design and Validation. *IEEE Transactions on Visualization and Computer Graphics* 15, 6 (2009), 921–928. doi:10.1109/TVCG.2009.111

[34] Tamara Munzner. 2014. *Visualization Analysis and Design*. AK Peters Visualization Series.

[35] C. L. Musatti. 1975. Stereokinetic phenomena and their interpretation. *Studies in Perception: Festschrift for Fabio Metelli. Martello-Giunti, Milano* (1975).

[36] Mark Nawrot and Keith Stroyan. 2012. Integration time for the perception of depth from motion parallax. *Vision research* 59 (2012), 64–71.

[37] Nami Ogawa, Takuji Narumi, and Michitaka Hirose. 2017. Swinging 3D lamps: A projection technique to convert a static 2D picture to 3D using wiggle stereoscopy. In *ACM SIGGRAPH'17 Posters*. 1–2.

[38] Hiroshi Ono, Brian J. Rogers, Masao Ohmi, and Mika E. Ono. 1988. Dynamic occlusion and motion parallax in depth perception. *Perception* 17, 2 (1988), 255–266.

[39] Michael Ortega and Wolfgang Stuerzlinger. 2018. Pointing at Wiggle 3D Displays. In *Proc. IEEE VR'18*. 335–340. doi:10.1109/VR.2018.8447552

[40] Carole Plasson, Renaud Blanch, and Laurence Nigay. 2022. Selection Techniques for 3D Extended Desktop Workstation with AR HMD. In *Proc. IEEE ISMAR'22*. 460–469. doi:10.1109/ISMAR55827.2022.00062

[41] Dennis R. Proffitt, Irvin Rock, Heiko Hecht, and Jim Schubert. 1992. Stereokinetic effect and its relation to the kinetic depth effect. *Journal of Experimental Psychology: Human Perception and Performance* 18, 1 (1992), 3.

[42] Jenny Reinhard. 2020. *Participants, Incentives, and User Studies: A Survey of CHI 2019*. Master's thesis. RWTH Aachen University.

[43] Jun Rekimoto and Mark Green. 1993. The information cube: Using transparency in 3D information visualization. In *Proc. WITS'93*, Vol. 13. 125–132.

[44] W. Richards. 1970. Stereopsis and stereoblindness. *Exp Brain Res.* 10, 4 (1970), 380–388. doi:10.1007/BF02324765

[45] George G. Robertson, Jock D. Mackinlay, and Stuart K. Card. 1991. Cone trees: animated 3D visualizations of hierarchical information. In *Proc. ACM CHI'91*. 189–194.

[46] Brian Rogers and Maureen Graham. 1979. Motion Parallax as an Independent Cue for Depth Perception. *Perception* 8, 2 (1979), 125–134.

[47] Brian Rogers and Maureen Graham. 1982. Similarities between motion parallax and stereopsis in human depth perception. *Vision research* 22, 2 (1982), 261–270.

[48] Bahador Saket, Arjun Srinivasan, Eric D Ragan, and Alex Endert. 2017. Evaluating interactive graphical encodings for data visualization. *IEEE transactions on visualization and computer graphics* 24, 3 (2017), 1316–1330.

[49] Stephen Smart and Danielle Albers Szafir. 2019. Measuring the separability of shape, size, and color in scatterplots. In *Proc. ACM CHI'19*. 1–14.

[50] George Sperling, Michael S. Landy, Barbara A. Dosher, and Mark E. Perkins. 1989. Kinetic depth effect and identification of shape. *J. Exp. Psychol.: Human Perception and Performance* 15, 4 (1989), 826.

[51] S. S. Stevens. 1957. On the psychophysical law. *Psychological Review* 64, 3 (1957), 153–181.

[52] S. S. Stevens. 1975. *Psychophysics: Introduction to its perceptual, neural and social prospects*. John Wiley & Sons.

[53] Danielle Albers Szafir, Steve Haroz, Michael Gleicher, and Steven Franconeri. 2016. Four types of ensemble coding in data visualizations. *Journal of vision* 16, 5 (2016), 11–11.

[54] C. Tominski, P. Schulze-Wollgast, and H. Schumann. 2005. 3D information visualization for time dependent data on maps. In *Proc. IV'05*. 175–181. doi:10.1109/IV.2005.3

[55] Stefan Treue, Masud Husain, and Richard A. Andersen. 1991. Human perception of structure from motion. *Vision Research* 31, 1 (1991), 59–75.

[56] Hermann von Helmholtz. 1925. *Helmholtz's Treatise on Physiological Optics*. Vol. 3. Optical Society of America.

[57] Mark Wagner. 2006. *The Geometries of Visual Space*. Psychology Press.

[58] Hans Wallach and Eileen B. Karsh. 1963. The modification of stereoscopic depth-perception and the kinetic depth-effect. *Am. J. Psychol.* 76, 3 (1963), 429–435.

[59] Hans Wallach and D. N. O'Connell. 1953. The kinetic depth effect. *J. of Exp. Psychol.* 45, 4 (1953), 205.

[60] Colin Ware. 2004. *Information visualization: perception for design* (1 ed.). Morgan Kaufmann.

[61] R. L. West and L. M. Ward. 1994. Constrained scaling. In *Proc 10th annual meeting of the International Society for Psychophysics*. 225–230.

[62] R. L. West, L. M. Ward, and R. Khosla. 2000. Constrained scaling: The effect of learned psychophysical scales on idiosyncratic response bias. *Perception & Psychophysics* 62 (2000), 137–151. doi:10.3758/BF03212067

[63] Daniel Wigdor, Chia Shen, Clifton Forlines, and Ravin Balakrishnan. 2007. Perception of elementary graphical elements in tabletop and multi-surface environments. In *Proc. ACM CHI'07* (San Jose, California, USA). 473–482. doi:10.1145/1240624.1240701

[64] L. M. Wilcox and J. M. Harris. 2010. *Fundamentals of Stereopsis*. 164–171. doi:10.1016/B978-0-12-374203-2.00237-2

[65] Wesley Willett, Bernhard Jenny, Tobias Isenberg, and Pierre Dragicevic. 2015. Lightweight Relief Shearing for Enhanced Terrain Perception on Interactive Maps. In *Proc. ACM CHI'15*. 3563–3572.

[66] Albert Yonas, Lincoln G. Craton, and William B. Thompson. 1987. Relative motion: Kinetic information for the order of depth at an edge. *Perception & Psychophysics* 41, 1 (1987), 53–59.

[67] Ahmad Yoonessi and Curtis L. Baker. 2011. Contribution of motion parallax to segmentation and depth perception. *Journal of vision* 11, 9 (2011), 13–13.

[68] Ahmad Yoonessi and Curtis L. Baker. 2013. Depth perception from dynamic occlusion in motion parallax: Roles of expansion-compression versus accretion-deletion. *J. Vis.* 13, 12 (2013), 10–10.