

AGri: Adaptive Thumbnails For Grid-based Visualizations

Steffen Frey*
University of Groningen

ABSTRACT

This work introduces AGri (Adaptive Thumbnails For Grid-based Visualizations), a method for dynamically adjusting thumbnails of spatiotemporal data—such as videos—to varying screen footprints in grid-based layouts. AGri aims to maximize thumbnail expressiveness, which quantifies how well similarity relationships among data members (e.g., video frames) are preserved. Thumbnails are generated via cropping, with crop windows optimized based on cumulative salience images. By modeling the trade-off between expressiveness and footprint size, AGri defines a curve—the AGri curve—representing Pareto-optimal visual representations. This curve enables dynamic selection of thumbnails suited to different grid sizes and resolutions. The approach is demonstrated on two datasets: a spatiotemporal ensemble from scientific experiments and an animated short film.

Index Terms: Grid-based Visualization, Thumbnails, Spatiotemporal Data, Video Visualization

1 INTRODUCTION

Videos are a form of spatiotemporal data that is highly prevalent in our daily lives and across various scientific domains [1, 5]. To support different use cases, dedicated visualization approaches have been proposed in fields such as medicine [31], sports [2, 25, 26, 32], and social media analytics [18]. A fundamental, general task across domains is to provide overviews that summarize video content [1, 5]. Storyboards are a common, domain-agnostic solution for this: they display (uniformly) sized video frame representations in a regular grid, typically reflecting the temporal sequence [5]. Traditionally, storyboards are shown at low grid resolutions (e.g., 3×4 images per page [3]).

Based on expert interviews, Afzal et al. [1] identified large-scale data visualization and pattern analysis as major requirements for video visualization. High-resolution grids are promising not only for presenting large datasets—where thumbnails represent video frames in grid cells—but also for employing layouts that reflect relationships between frames to support pattern analysis (e.g., revealing clusters and outliers). Dimensionality reduction (DR) techniques can capture these relationships by projecting video frames into 2D scatterplots, such that similarity relations between frames—i.e., the underlying data structure—are preserved [6]. Overlap-removal methods generate grids from such plots by rearranging thumbnails while aiming to preserve the layout characteristics of the original DR projection [7, 8, 10, 11, 12, 14, 15, 17, 22].

There is a fundamental trade-off in grid-based visualization [10]: as grid resolution increases, a greater number of thumbnails can be shown, allowing for more comprehensive coverage of the video. However, the footprint of each thumbnail—the space it occupies on the screen—decreases, which hampers expressiveness, defined here as the degree to which thumbnails preserve similarity relationships between video frames. This notion of expressiveness is

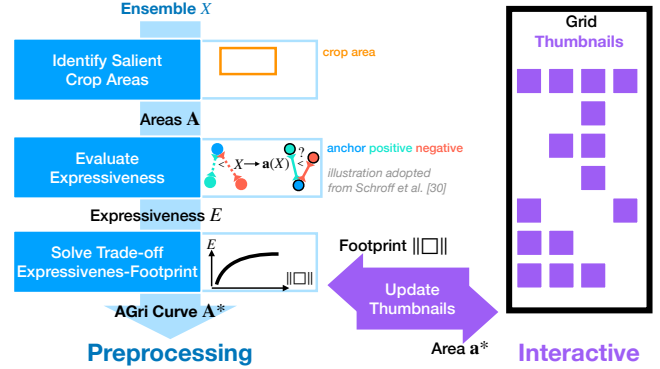


Figure 1: AGri determines the AGri curve during preprocessing, which defines optimal thumbnail representations across different footprints. At runtime, AGri dynamically selects thumbnails from this curve as the grid view—and thus the footprint—changes.

aligned with the dimensionality reduction (DR) objective of preserving the inherent structure of spatiotemporal data, as discussed above. In the visualization literature, the challenge of “show[ing] a larger amount of data in a readable way” [27] is often referred to as *visual scalability* [4, 9, 28]. In grid-based representations, thumbnail footprints can also change dynamically in interactive scenarios, such as when adjusting the grid resolution and thereby the number of thumbnails shown [7, 10, 17].

In this paper, we introduce AGri, a method to dynamically adapt thumbnail representations to different footprints, optimizing the trade-off between expressiveness and footprint for grid-based visualizations of spatiotemporal data such as videos. AGri generates thumbnails via cropping, allowing the display to focus on the most relevant regions of the original video frames. The position of each crop window is optimized using cumulative salience images. From a range of crop windows of different sizes, the Pareto-optimal AGri curve enables the dynamic selection of thumbnail representations based on their footprint within a grid. In doing so, AGri offers a novel approach to the visual scalability challenge of presenting spatiotemporal data through thumbnails in grid-based layouts. We demonstrate AGri on two use cases: (i) *droplet*, a 2D spatiotemporal ensemble from scientific experiments [13], and (ii) *Big Buck Bunny (BBB)*, an animated short film featuring forest animals [29].

2 METHOD

Given a 2D spatiotemporal ensemble—a video (collection)— X , AGri aims to determine the AGri curve, a set of thumbnail representations for grid-based visualizations that are optimal with respect to the trade-off between thumbnail expressiveness E and pixel footprint $||\square||$. The AGri curve enables interactive, adaptive selection of thumbnail representations suitable for different grid sizes and resolutions. The respective steps are outlined below (also see Fig. 1).

Identify Salient Crop Areas A

From the 2D pixel resolution of ensemble members (video frames), we extract different areas (2D crop window extents) A that will be

*e-mail: s.d.frey@rug.nl

Input: areas \mathbf{A} , expressiveness E

Result: AGri curve parameters \mathbf{A}^*

```

1  $\mathbf{A} \leftarrow \text{sort}(\mathbf{A}, \text{key} = E(\cdot), \text{order} = \text{descending});$ 
2  $\mathbf{A}^* \leftarrow \text{front}(\mathbf{A});$ 
3 forall  $\mathbf{a} \in \mathbf{A}$  do
4   if  $\|\mathbf{a}\| \leq \|\text{back}(\mathbf{A}^*)\|$  then
5      $\mathbf{A}^* \leftarrow \mathbf{A}^* \cup \{\mathbf{a}\};$ 
6 return  $\mathbf{A}^*;$ 

```

Algorithm 1: Identification of the AGri curve.

considered throughout the AGri procedure. We construct this set iteratively: starting from the full resolution \mathbf{a}_0 , we incrementally reduce the number of pixels in the x -direction, adjusting y to maintain the original aspect ratio as closely as possible. A new configuration \mathbf{a}_1 is selected when the total number of footprint pixels $\|\mathbf{a}\|$ differs by a certain factor ($\approx 10\%$ in this work). This process is repeated, where each \mathbf{a}_{i+1} is derived from \mathbf{a}_i , until \mathbf{a}_i contains four pixels or fewer.

To generate a thumbnail $\mathbf{a}(x)$ with area $\mathbf{a} \in \mathbf{A}$ for a member $x \in X$, we need to determine a specific crop window offset, as the crop can be positioned freely. To find the best offset, we use a simple salience measure for spatiotemporal data. We compute difference images with both the previous and next frames, and select the one with the lower total difference. From this, we derive a cumulative salience image, reducing each evaluation to four lookups. We then evaluate all possible offsets (with a step size of 4 pixels in this work) and choose the one with the highest salience sum.

Evaluate Expressiveness E

We quantify expressiveness E using triplet loss, a machine learning approach that compares relative distances and is effective even with few samples [30]. Adopting triplet loss, AGri evaluates whether similarity relationships among ensemble members—video frames— $x_0, x_1, x_2 \in X$ are preserved in their respective thumbnails $\mathbf{a}(x_0), \mathbf{a}(x_1), \mathbf{a}(x_2)$. To do this, we randomly sample ensemble members to form a fixed set of triplets T , where each triplet $(x_a, x_p, x_n) \in T$ consists of an anchor x_a , a positive x_p , and a negative x_n , such that:

$$\Delta_{\text{MSE}}(x_a, x_p) < \Delta_{\text{MSE}}(x_a, x_n),$$

with Δ_{MSE} denoting mean squared error (considering $\|T\| = 1024$ triplets in this work). We then assess whether the same relationship holds for their thumbnails, using the perceptual similarity metric LPIPS [34]. Expressiveness $E(\mathbf{a}) \in [0, 1]$ for $\mathbf{a} \in \mathbf{A}$ is defined as:

$$E(\mathbf{a}) := \frac{1}{\|T\|} \sum_{(x_a, x_p, x_n) \in T} \begin{cases} 1, & \text{if } \Delta_{\text{LPIPS}}(\mathbf{a}(x_a), \mathbf{a}(x_p)) < \Delta_{\text{LPIPS}}(\mathbf{a}(x_a), \mathbf{a}(x_n)) \\ 0, & \text{otherwise.} \end{cases}$$

Solve Trade-Off: Expressiveness vs. Footprint

We now compute the AGri curve, which contains the Pareto-optimal area subset $\mathbf{A}^* \subseteq \mathbf{A}$, considering both area expressiveness $E(\mathbf{A})$ and footprint $\|\mathbf{a}\|$. Our algorithm (Alg. 1) begins by sorting the design space \mathbf{A} in descending order of expressiveness (l. 1). The configuration with the highest expressiveness is added first to the AGri curve (l. 2). Then, for each $\mathbf{a} \in \mathbf{A}$, if its footprint $\|\mathbf{a}\|$ is smaller than that of the last added element (l. 4), we include it in the AGri curve (l. 5)—thus maintaining Pareto-optimality with respect to decreasing size.

Update Thumbnails

The AGri curve enables efficient runtime selection of a crop area \mathbf{a}^* for a given requested footprint $\|\square\|$. We choose the configuration

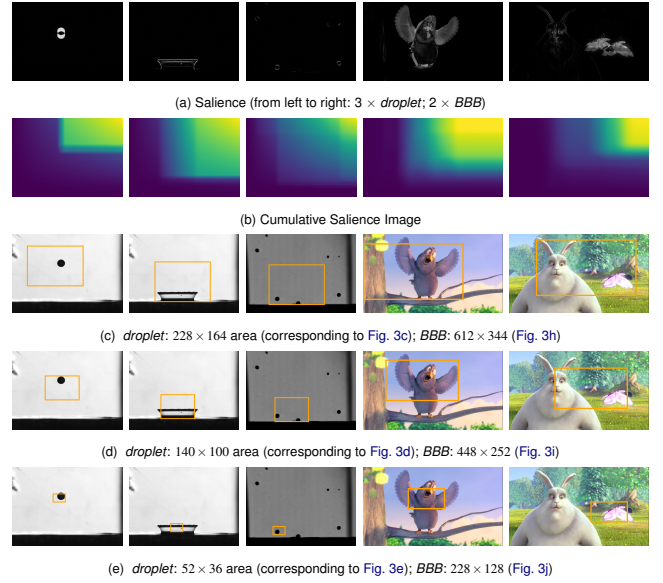


Figure 2: Saliency-based identification of crop offsets. (a) Saliency computed from differences to temporally adjacent frames. (b) Cumulative saliency image enabling efficient evaluation of saliency within arbitrary rectangular regions. (c-e) Selected crop areas for varying footprint sizes.

with the largest footprint smaller than $\|\square\|$ on the AGri curve \mathbf{A}^* :

$$\mathbf{a}^* \leftarrow \arg \max_{\mathbf{a} \in \mathbf{A}^*} E(\mathbf{a}) \quad \text{s.t. } \|\mathbf{a}\| \leq \|\square\|.$$

3 CASE STUDIES

We consider two datasets: (i) *droplet*—an experimental dataset analyzing the impact of a droplet hitting a thin fluid film [13], and (ii) *BBB*—the animated comedy short film Big Buck Bunny, featuring forest animals [29]. The *droplet* dataset contains 6120 experimental grayscale images at a resolution of 448px \times 320px. *BBB* consists of 1431 RGB video frames, sampled every tenth frame from the original video, at a resolution of 420p (853px \times 480px). AGri is implemented in Python 3.13 using NumPy, and employs the *lpips* module to assess perceptual differences [34]. Since different crop areas can be evaluated independently, we use a parallel CPU implementation. On an entry-level MacBook Air (M4), the complete AGri pipeline takes approximately 30 min for *droplet* and 5 h for *BBB*.

Salient Crop Areas (Fig. 2)

(a) The saliency measure shows to be expressive for both *droplet* and *BBB*. For *droplet*, the focus lies on the falling and deforming droplet, the expanding crown, and the resulting splashes. In *BBB*, the emphasis is on the most active characters in the given frame. (b) presents the cumulative saliency images, which are aggregated from individual saliency fields to enable fast evaluation of saliency within arbitrary regions. (c-e) show selected crop areas for different footprints—configurations along the AGri curve (see Fig. 3 and Fig. 4 below). These areas generally align well with regions of high saliency, but a key limitation becomes apparent: fixed-size crop windows cannot always adapt to the full extent or shape of saliency distributions. For example, in (c), the crop area in *droplet* is larger than the small salient region caused by the falling droplet, yet this size is necessary to fully encompass the expanding crown. In *BBB*, the same window is oversized for the bird but appropriate when both the bunny and the butterfly are present. In (e), a smaller crop area is sufficient to capture the droplet or butterfly, but

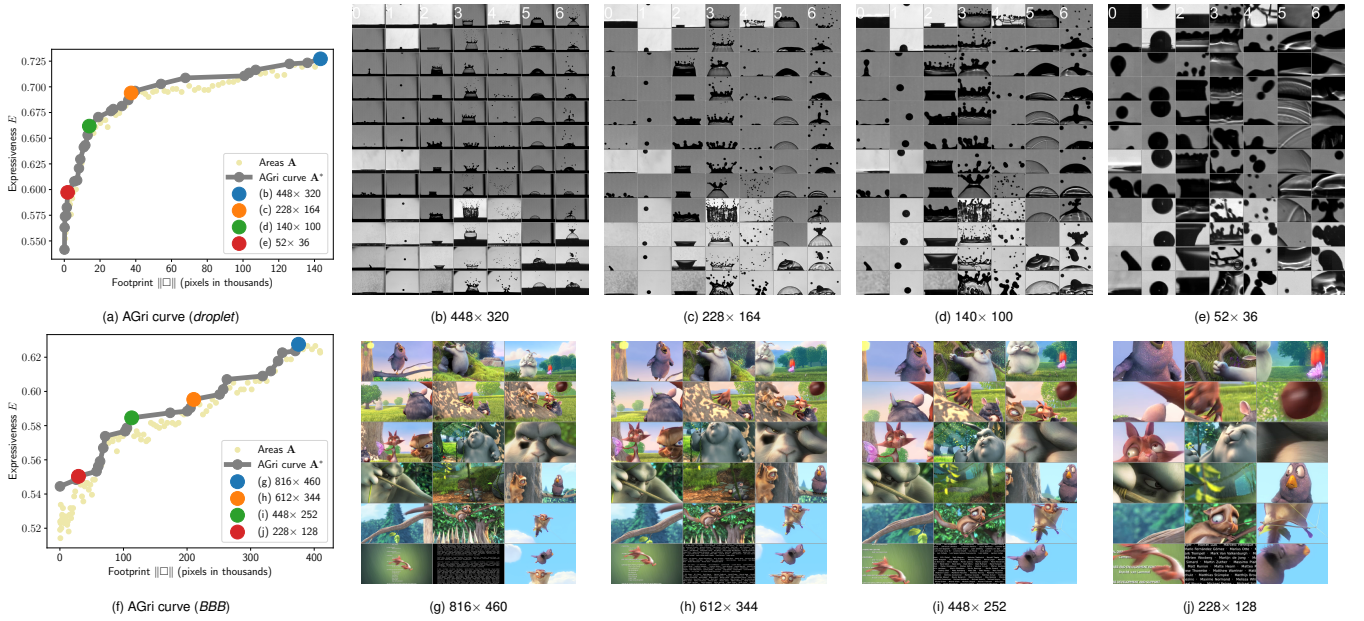


Figure 3: (a & f) The AGri curve depicts Pareto-optimal solutions A^* (gray) for *droplet* and *BBB*, respectively; other areas ($A - A^*$) are shown in yellow for reference. (b-e & g-j) The resulting thumbnails for selected areas (colored in (a & f), respectively) are presented in a small grid.

may omit other contextually relevant parts. Note that at this smaller scale, the step size of 4 pixels used for evaluating potential offsets plays a more significant role. In both over- and under-coverage scenarios, the optimal crop positioning can depend on subtle factors. Nevertheless, the selected crops retain strong expressiveness, and the consistent preservation of spatial scale remains a key advantage of the design.

AGri curve (Fig. 3)

The AGri curve identifies thumbnail configurations that achieve reduced footprint with comparatively small loss in expressiveness—shown in (a) for *droplet* and in (f) for *BBB*. The shape of the AGri curve for both datasets indicates that the trade-off between footprint and expressiveness is nonlinear. The annotated points $A - A^*$ (in yellow) illustrate that increasing the footprint does not necessarily lead to increased expressiveness. For *droplet*, images are arranged based on expert-assigned labels in (b-e) (as indicated by the numbers at the top of the columns, which represent different behavior types [13]). *BBB* uses a temporal grid layout akin to a storyboard in (g-j). (b) For *droplet*, the highest expressiveness is achieved using the full image. (g) For *BBB*, near-maximal expressiveness is obtained with a crop area covering 92 % of the frame. This highlights the influence of using different metrics: while Δ_{MSE} is used to define similarity in the triplet sampling, Δ_{LPIPS} is employed to evaluate thumbnail similarity. As a result, even the reference configuration does not achieve perfect expressiveness ($E = 1$).

(c) For *droplet*, a footprint reduction to $\approx 25\%$ still maintains high expressiveness, with thumbnails conveying the core dynamics. (d) Moving further along the “elbow” of the AGri curve, high expressiveness is retained even at $\approx 10\%$ of the original footprint. While the main features are still captured, large structures (e.g., extensive crowns) may be only partially visible. (e) Reducing the footprint further—to just over 1 %—results in a noticeable drop in expressiveness E . Though small features like individual droplets are still captured, distinguishing between them and more complex events (e.g., splashes) becomes difficult. Large-scale features, such as developing crowns, are often only partially visible. That said, key visual characteristics remain preserved in the crop images.

(h) For *BBB*, a significant footprint reduction to $\approx 51\%$ also leads to only a small loss in expressiveness, although the trade-off is less favorable than for *droplet*. We attribute this to the higher visual complexity of *BBB* frames. (i) Continuing along the AGri curve, expressiveness remains high at $\approx 27\%$ footprint, with thumbnails still capturing the main activity in most scenes. (j) Reducing the footprint to $\approx 7\%$ leads to further loss in expressiveness, but thumbnails still capture the dominant actions. Some elements, however, may be omitted—e.g., in Fig. 2e (right), the butterfly is preserved, while the bunny is no longer visible.

Thumbnail Update (Fig. 4)

We now apply the AGri curve—specifically the configurations presented earlier in Fig. 3—to grid-based visualizations. To this end, we use Sca²Gri (Scalable Gridified Scatterplots), a recently introduced post-processing technique for generating grid-based layouts, utilizing the provided implementation [10]. Sca²Gri produces grids with varying numbers of cells which is controlled by modifying the horizontal resolution (g_x). Starting from a 2D embedding of the ensemble generated via UMAP [21], Sca²Gri assigns (a subset of) ensemble members to grid cells such that a certain assignment distance is not exceeded (only one member can be presented per cell, members that cannot be assigned under the distance constraint are discarded). As (g_x increases, the grid can show more thumbnails, but to maintain constant overall resolution, the footprint $\mathbf{a}_{\|\square\|}$ of each thumbnail decreases accordingly.

For *droplet*, (a) shows a low-resolution grid with $g_x = 5$, using the AGri configuration from Fig. 3b. This setting preserves the full detail of each thumbnail, but only a small portion of the ensemble is displayed. Increasing the resolution to $g_x = 10$ in (b) allows a broader selection of ensemble members to be shown, while the thumbnails still convey the core features of each frame. With $g_x = 16$ in (c), the grid presents a comprehensive view of the structural variability in the ensemble, while retaining good visual expressiveness in the thumbnails. This configuration provides a balanced trade-off between grid resolution g_x and thumbnail footprint $\mathbf{a}_{\|\square\|}$. Further increasing the resolution to $g_x = 43$ in (d) significantly expands coverage of the ensemble, but the thumbnail footprints be-

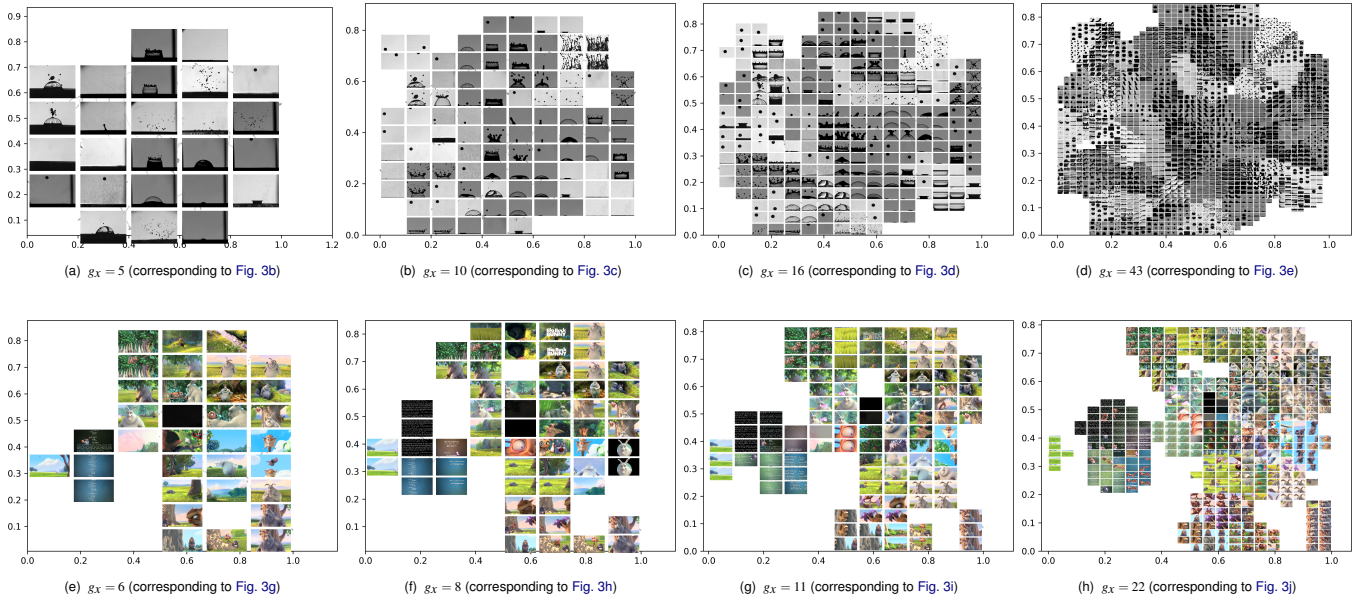


Figure 4: Grids created via Sca²Gri [10] with thumbnails from Fig. 3. Assuming a fixed total resolution in the order of a full HD screen, the requested thumbnail footprint $\|\square\|$ is accordingly decreased when increasing (horizontal) grid resolution g_x .

come small, reducing their expressiveness—as discussed above by means of Fig. 3e.

For *BBB*, (e) illustrates a configuration with ($g_x = 6$) where the thumbnails are highly expressive, but only a limited number of frames can be shown. Increasing the resolution to $g_x = 8$ in (f) enables a greater variety of frames to be displayed, with thumbnails still capturing the most salient elements. At $g_x = 11$, shown in (g), the grid more accurately reflects the structure of the underlying 2D embedding while still maintaining expressive thumbnails. Finally, (h) presents a grid with $g_x = 22$, which best preserves the layout derived from dimensionality reduction. However, the thumbnails are significantly smaller, leading to lower expressiveness, as previously discussed in the context of Fig. 3j.

4 DISCUSSION AND FUTURE WORK

We now discuss promising directions for extending, improving, and generalizing AGri in future work.

Perceptual Aspects. In the grid visualizations considered above, we focused primarily on technical aspects and did not explicitly account for perceptual factors. Healey and Sawant conducted psychophysical experiments to investigate the limits of resolution and visual angle in visualization [16]. They found that a minimum number of pixels (resolution) and subtended physical area on the retina (visual angle) are required to distinguish visual features, and that these limits can easily be exceeded in typical visualizations. This makes visual angle a key factor in characterizing the properties of the eye and early vision (see Ware [33]). Incorporating these perceptual considerations into AGri’s thumbnail adaptation process would be a valuable extension, especially in supporting the scaling of visualizations to a range of display scenarios—from small devices like smartwatches to large-scale displays like powerwalls.

Heterogeneous Areas. Area selection can produce multiple similarly optimal solutions when the target footprint is either very small or very large. One possible direction to address this challenge is to allow heterogeneous (variable-sized) crop areas that better align with the extent of salience in each frame. To maintain a consistent visual footprint, selected areas could then be down-sampled accordingly to fit the required display size. However, this comes at the cost of losing the benefit of uniform thumbnail sizes

in grid visualizations, which helps preserve size-related interpretations of visual features.

Expressiveness Quantification. AGri currently quantifies expressiveness using a triplet-loss-inspired approach from machine learning [30], where triplets are constructed based on data similarity (Δ_{MSE}). A key strength of this approach is that it requires only the raw data as input—no labels or metadata are needed. However, in some use cases, additional information is available that could further guide expressiveness modeling. For example, in *droplet*, expert-assigned labels (corresponding to class numbers in Fig. 3b–e) could be used to form semantically meaningful triplets, where *anchor* and *positive* sample come from the same class, and the *negative* sample from a different one. The improvement in expressiveness could further be underlined and/or driven by empirical studies.

Generalization. We demonstrated AGri for two different datasets: *droplet* and *BBB*. In both cases, we deal with temporal data, and salience is computed using simple frame-difference measures. More general salience models could be considered in the future to support other types of data and further improve expressiveness. In addition, extending the parameter space \mathbf{A} (i.e., the space of candidate crop areas in the current work) could serve as a foundation for more general thumbnail design strategies (e.g., complementing existing models of visualization design [19, 20, 23, 24]). Computationally, this would further induce the need for efficient search approaches for high-dimensional visualization design spaces \mathbf{A} , as well as implementations for GPUs and/or distributed environments (exploiting the massive parallelism induced by large $\|\mathbf{A}\|$).

5 CONCLUSION

We presented AGri, a novel method for adaptive thumbnail generation in grid-based visualizations, designed to maximize expressiveness for a range of thumbnail footprints. AGri uses cumulative salience images to guide the selection of crop areas, and computes the AGri curve to balance the trade-off between expressiveness and footprint. This enables flexible, data-aware adaptations of thumbnails while preserving similarity relationships—i.e., the structure—of the underlying spatiotemporal ensemble or video. In doing so, AGri contributes an approach to enhance both the scalability and interpretability of thumbnails in grid-based visualizations.

ACKNOWLEDGMENTS

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project Number 327154368 - SFB 1313.

REFERENCES

- [1] S. Afzal, S. Ghani, M. M. Hittawe, S. F. Rashid, O. M. Knio, M. Hadwiger, and I. Hoteit. Visualization and visual analytics approaches for image and video datasets: A survey. *ACM Trans. Interact. Intell. Syst.*, 13(1), 2023. doi: 10.1145/3576935 1
- [2] G. Andrienko, N. Andrienko, G. Anzer, P. Bauer, G. Budziak, G. Fuchs, D. Hecker, H. Weber, and S. Wrobel. Constructing spaces and times for tactical analysis in football. *IEEE Transactions on Visualization and Computer Graphics*, 27(4):2280–2297, 2021. doi: 10.1109/TVCG.2019.2952129 1
- [3] W. Bailer and G. Thallinger. A framework for multimedia content abstraction and its application to rushes exploration. In *Proceedings of the 6th ACM International Conference on Image and Video Retrieval, CIVR '07*, p. 146–153. ACM, New York, 2007. doi: 10.1145/1282280.1282307 1
- [4] O. Bieh-Zimmert and C. Felden. Shaping unlimited patterns: A vision for state-of-the-art visual scalability. In *Proceedings of the 6th International Conference on Management of Emergent Digital EcoSystems, MEDES '14*, p. 72–77. ACM, New York, 2014. doi: 10.1145/2668260.2668279 1
- [5] R. Borgo, M. Chen, B. Daubney, E. Grundy, G. Heidemann, B. Höferlin, M. Höferlin, H. Lettice, D. Weiskopf, and X. Xie. State of the art report on video-based graphics and video visualization. *Computer Graphics Forum*, 31(8):2450–2477, 2012. doi: 10.1111/j.1467-8659.2012.03158.x 1
- [6] D. B. Coimbra, R. M. Martins, T. T. Neves, A. C. Telea, and F. V. Paulovich. Explaining three-dimensional dimensionality reduction plots. *Information Visualization*, 15(2):154–172, 2016. doi: 10.1177/1473871615600010 1
- [7] R. Cutura, C. Morariu, Z. Cheng, Y. Wang, D. Weiskopf, and M. Sedlmair. Hagrid: using hilbert and gopher curves to gridify scatterplots. *Journal of Visualization*, 25(6):1291–1307, 2022. doi: 10.1145/3481549.3481569 1
- [8] T. Dwyer, K. Marriott, and P. J. Stuckey. Fast node overlap removal. In P. Healy and N. S. Nikolov, eds., *Graph Drawing*, pp. 153–164. Springer, Berlin, Heidelberg, 2006. doi: 10.1007/11618058_15 1
- [9] S. G. Eick and A. F. Karr. Visual scalability. *Journal of Computational and Graphical Statistics*, 11(1):22–43, 2002. doi: 10.1198/106186002317375604 1
- [10] S. Frey. Sca2Gri: Scalable gridified scatterplots. *Computer Graphics Forum*, 44(3), 2025. doi: 10.1111/cgf.70141 1, 3, 4
- [11] E. Gansner and Y. Hu. Efficient, proximity-preserving node overlap removal. *J. Graph Algorithms Appl.*, 14:53–74, 2010. doi: 10.7155/jgaa.00198 1
- [12] M. v. Garderen, B. Pampel, A. Nocaj, and U. Brandes. Minimum-Displacement Overlap Removal for Geo-referenced Data Visualization. *Computer Graphics Forum*, 2017. doi: 10.1111/cgf.13199 1
- [13] A. Geppert, D. Chatzianagnostou, C. Meister, H. Goma, G. Lamanna, and B. Weigand. Classification of Impact Morphology and Splashing/Deposition Limit For N-Hexadecane. *Atomization and Sprays*, 26(10), 2016. doi: 10.1615/AtomizSpr.2015013352 1, 2, 3
- [14] E. Gomez-Nieto, F. S. Roman, P. Pagliosa, W. Casaca, E. S. Helou, M. C. F. de Oliveira, and L. G. Nonato. Similarity preserving snippet-based visualization of web search results. *IEEE Transactions on Visualization and Computer Graphics*, 20(3):457–470, 2014. doi: 10.1109/TVCG.2013.242 1
- [15] K. Hayashi, M. Inoue, T. Masuzawa, and H. Fujiwara. A layout adjustment problem for disjoint rectangles preserving orthogonal order. In *Graph Drawing*, 1998. doi: 10.1007/3-540-37623-2_14 1
- [16] C. G. Healey and A. P. Sawant. On the limits of resolution and visual angle in visualization. *ACM Transactions on Applied Perception*, 9(4):20:1–20:21, Oct. 2012. doi: 10.1145/2355598.2355603 4
- [17] G. M. Hilaraca, W. E. Marcilio-Jr, D. M. Eler, R. M. Martins, and F. V. Paulovich. A grid-based method for removing overlaps of dimensionality reduction scatterplot layouts. *IEEE Transactions on Visualization and Computer Graphics*, 30(8):5733–5749, 2024. doi: 10.1109/TVCG.2023.3309941 1
- [18] M. Itoh, M. Toyoda, T. Kamijo, and M. Kitsuregawa. Visualizing flows of images in social media. In *2012 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp. 229–230, 2012. doi: 10.1109/VAST.2012.6400539 1
- [19] G. Kindlmann and C. Scheidegger. An Algebraic Process for Visualization Design. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2181–2190, 2014. doi: 10.1109/TVCG.2014.2346325 4
- [20] J. Mackinlay, P. Hanrahan, and C. Stolte. Show Me: Automatic Presentation for Visual Analysis. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1137–1144, 2007. doi: 10.1109/TVCG.2007.70594 4
- [21] L. McInnes, J. Healy, N. Saul, and L. Grossberger. UMAP: Uniform manifold approximation and projection. *The Journal of Open Source Software*, 3(29):861, 2018. doi: 10.21105/joss.00861 3
- [22] K. Misue, P. Eades, W. Lai, and K. Sugiyama. Layout adjustment and the mental map. *J. Vis. Lang. Comput.*, 6:183–210, 1995. doi: 10.1006/jvlc.1995.1010 1
- [23] V. O. Mittal, G. Carenini, J. D. Moore, and S. Roth. Describing complex charts in natural language: a caption generation system. *Comput. Linguist.*, 24(3):431–467, Sept. 1998. 4
- [24] D. Moritz, C. Wang, G. L. Nelson, H. Lin, A. M. Smith, B. Howe, and J. Heer. Formalizing visualization design knowledge as constraints: Actionable and extensible models in draco. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):438–448, 2019. doi: 10.1109/TVCG.2018.2865240 4
- [25] C. Perin, R. Vuilleumot, and J.-D. Fekete. Soccerstories: A kick-off for visual soccer analysis. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2506–2515, 2013. doi: 10.1109/TVCG.2013.192 1
- [26] T. Polk, D. Jäckle, J. Häubler, and J. Yang. Courttime: Generating actionable insights into tennis matches using visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):397–406, 2020. doi: 10.1109/TVCG.2019.2934243 1
- [27] G. Richer, A. Pister, M. Abdelaal, J.-D. Fekete, M. Sedlmair, and D. Weiskopf. Scalability in visualization. *IEEE Transactions on Visualization and Computer Graphics*, 30(7):3314–3330, 2024. doi: 10.1109/TVCG.2022.3231230 1
- [28] G. Robertson, D. Ebert, S. Eick, D. Keim, and K. Joy. Scale and complexity in visual analytics. *Information Visualization*, 8(4):247–253, 2009. doi: 10.1057/ivs.2009.23 1
- [29] T. Roosendaal. Big buck bunny. In *ACM SIGGRAPH ASIA 2008 Computer Animation Festival, SIGGRAPH Asia '08*, p. 62. ACM, New York, 2008. doi: 10.1145/1504271.1504321 1, 2
- [30] F. Schroff, D. Kalenichenko, and J. Philbin. Facenet: A unified embedding for face recognition and clustering. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015. doi: 10.1109/cvpr.2015.7298682 2, 4
- [31] H. Song, J. Lee, T. J. Kim, K. H. Lee, B. Kim, and J. Seo. Gazedx: Interactive visual analytics framework for comparative gaze analysis with volumetric medical images. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):311–320, 2017. doi: 10.1109/TVCG.2016.2598796 1
- [32] M. Stein, H. Janetzko, A. Lamprecht, T. Breitkreutz, P. Zimmermann, B. Goldlücke, T. Schreck, G. Andrienko, M. Grossniklaus, and D. A. Keim. Bring it to the pitch: Combining video and movement data to enhance team sport analysis. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):13–22, 2018. doi: 10.1109/TVCG.2017.2745181 1
- [33] C. Ware. *Information Visualization: Perception for Design*. Elsevier, 2020. doi: 10.1016/C2016-0-02395-1 4
- [34] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 586–595, 2018. doi: 10.1109/CVPR.2018.00068 2