

# A Study of Interactive Decision-Making Techniques in Multimodal Big Data Visual Analytics

Yihao Ning\*

School of Converged Media Center, Hainan Vocational University of Science and Technology, Haikou, Hainan, 571126, China.

\*Corresponding author: 124027287@student.newinti.edu.my

**Abstract:** *With the rapid advancement of big data technology, processing and analyzing multimodal, heterogeneous data have become a central challenge in the field of decision science. Multimodal big data visual analytics aims to transform complex data into actionable decision insights by integrating visual interaction with computational models. This study focuses on interactive decision-making techniques in this domain, systematically exploring their theoretical foundations and technical architecture. The research first analyzes the core challenges associated with the representation learning and heterogeneous fusion of multimodal data, the cognitive theoretical basis of visual analytics, and the construction of decision task-driven analytical models. Subsequently, it constructs a technical framework for visual analytics oriented toward interactive decision-making, encompassing key technologies such as multi-scale dynamic visualization, interactive feature exploration, uncertainty quantification, and decision reasoning. Finally, it proposes design principles for decision-maker-centered system integration, methods for enhancing interactive intelligence, and a multidimensional performance evaluation framework. This study provides theoretical references and technical pathways for building high-performance interactive decision support systems.*

**Keywords:** *Multimodal big data; Visual analytics; Interactive decision-making; Representation learning; Uncertainty quantification*

## Introduction

Under the data-driven research paradigm, decision-makers are confronted with an increasingly complex data environment characterized by massive scale, diverse modalities, and dynamic evolution. Traditional data analysis methods often exhibit limitations in areas such as information integration, pattern discovery, and causal reasoning when processing such multimodal big data, making it difficult to support a deep understanding of complex problems and agile decision-making. Consequently, visual analytics, which integrates human cognitive strengths with machine computational power-particularly interactive decision-making techniques emphasizing human-machine collaboration-has become a key pathway to address this challenge. The significance and necessity of this study lie in systematically reviewing and deepening the theoretical foundations of interactive decision-making techniques within multimodal big data environments, and in constructing a complete technological chain spanning from data fusion and visual encoding to decision reasoning. This aims to bridge the semantic gap between raw data and high-level decision-making intelligence, thereby enhancing analysts' efficiency and judgment quality in information-overloaded environments. The subsequent sections of this paper will elaborate from three aspects: theoretical foundations, technological architecture, and system implementation, providing a structured reference for theoretical development and practical system building in related fields.

## 1. Theoretical Foundations and Core Challenges in Multimodal Big Data Visual Analytics

### 1.1 Representation Learning and Heterogeneous Fusion of Multimodal Data

Effective analysis of multimodal data begins with the unification and fusion of its underlying representations. The core objective of representation learning is to construct a shared semantic embedding space for heterogeneous modal data (such as text, images, temporal signals, and topological structures), ensuring that information from different sources but with semantic relevance is

geometrically proximate within this space. This approach typically involves deep neural network architectures, such as cross-modal encoders and attention mechanisms, designed to capture the complex nonlinear correspondences between modalities. Heterogeneous fusion further integrates these aligned representations into unified features usable for decision-making. The fusion strategy must address differences in data quality, signal-to-noise ratio, and update frequency across modalities. A key challenge lies in designing fusion mechanisms capable of dynamically weighing the contributions of each modality to avoid information redundancy or dominance, while ensuring the fused representations are interpretable and actionable for subsequent interactive analysis tasks.

### ***1.2 Cognitive Theory and Perceptual Enhancement in Visual Analytics***

The effectiveness of visual analytics is fundamentally constrained by the cognitive principles of human information processing; therefore, its design must be rooted in a solid foundation of cognitive science. Cognitive load theory demands that interaction design strive to optimize the occupation of working memory. By pursuing simplicity, consistency, and predictability in visual encoding, it aims to minimize users' unnecessary cognitive shifts and mental effort during information interpretation. Principles such as Gestalt principles, color perception theory, and visual saliency computation provide concrete guidance for view design, effectively directing users' attention to key patterns, correlations, and potential anomalies. Operating within this theoretical framework, perceptual enhancement techniques aim to transcend the passive presentation of static information. They intuitively reveal inherent data evolution trends and causal relationships by utilizing dynamic visual variables such as animation, traces, and morphing, or expand the effective bandwidth of human information perception through techniques like multi-view coordination, focus+context, and immersive visualization. The core challenge and frontier in this field lie in how to precisely quantify the aforementioned, relatively abstract cognitive and perceptual principles into computable and optimizable specific visual encoding parameters and interactive dynamics models. Furthermore, it is crucial to empirically validate the substantive improvement these enhancement methods bring to users' speed of insight formation, decision confidence, and analytical accuracy within complex multimodal decision-making contexts<sup>[1]</sup>.

### ***1.3 Construction of Decision Task-Driven Visual Analytics Models***

The effectiveness of interactive decision-making is closely dependent on a high degree of alignment between the analytical process and the inherent logical structure of the task. Therefore, constructing a formalized task model is necessary as a blueprint for system design. Decision task-driven visual analytics models aim to provide a structured analytical framework for specific types of decision problems, such as classification, root cause analysis, prediction, and solution comparison. This involves deconstructing high-level, ambiguous decision objectives in a top-down manner into sequences of fundamental analytical actions that can be interactively implemented, for example, hypothesis generation, evidence exploration, scenario evaluation, trade-off analysis, and conclusion confirmation. The model construction process requires the abstraction and formal definition of common, key elements of decision tasks. These elements include, but are not limited to, the uncertainty level of the problem, time constraints for the decision, the multiplicity and potential conflict of objectives, and stringent requirements for the interpretability of conclusions. Subsequently, these elements must be systematically mapped to the most suitable visual metaphors, visual encoding schemes, and sets of interaction primitives. The innovation lies in establishing a meta-task model that possesses both guiding and flexible properties. This model should be capable of providing effective analytical pathway guidance that conforms to cognitive principles, thereby preventing users from becoming lost in complex data. Simultaneously, it must fully respect and preserve users' autonomy and creativity to dynamically adjust their analytical focus, redefine problem boundaries, or even shift their exploration direction during the interaction process.

## **2. Visual Analytics Technical Framework for Interactive Decision-Making**

### ***2.1 Multi-Scale Dynamic Visualization and View Coordination Techniques***

#### ***2.1.1 Data Abstraction Mechanisms Based on Semantic Zooming and Level-of-Detail***

Traditional scale transformation primarily relies on geometric zooming. For multimodal semantic data, it is necessary to introduce semantic zooming techniques. This technique dynamically aggregates or expands visual elements based on data attributes, statistical features, or clustering relationships. In

this process, the level-of-detail model is bound to the semantic hierarchy of the data, ensuring a smooth and semantically coherent visual transition from an overview of topic distributions down to specific raw data points. This mechanism allows users to quickly locate subsets of data of interest without expending cognitive resources on irrelevant details.

### ***2.1.2 Dynamic Visual Encoding Strategies Supporting the Continuous Evolution of States***

Dynamic visualization refers not merely to the switching of views, but emphasizes the continuous, smooth change of visual element attributes to reflect the evolution of data states or models. This involves the mapping of time, parameter spaces, or model iteration steps, utilizing encoding methods such as animation, morphing, and trajectory overlay to reveal the intrinsic dynamics and causal flows within the data. The key technical challenge lies in maintaining visual object constancy throughout the animation process and designing transition effects that conform to perceptual principles, thereby assisting users in tracking changes and identifying anomalous evolutionary patterns.

### ***2.1.3 Coordinated Interaction and Context Maintenance in Multi-View Associations***

Within multi-view analysis environments, view coordination techniques are crucial for maintaining a global analytical context. By employing predefined or user-defined linkage rules (such as brushing, focusing, and linking), these techniques ensure that an interactive operation in one view automatically triggers visual feedback updates in other associated views. This coordination can extend beyond direct associations of data points to deeper linkages based on statistical summaries, feature similarity, or model prediction outcomes. The goal is to create an integrated visual analytics environment, enabling users to develop a holistic understanding of complex problems through the juxtaposition, comparison, and correlation of information from different perspectives<sup>[2]</sup>.

## ***2.2 Interactive Feature Exploration and Pattern Discovery Techniques***

### ***2.2.1 Active Feature Selection and Visual Manipulation Mechanisms***

Going beyond automated feature selection, this mechanism allows users to define and manipulate features through direct interaction with the visual interface. Users can employ brush tools to delineate regions of interest within scatter plots or projection maps, thereby defining new features based on visual clusters; alternatively, they can dynamically alter feature weights and combinations by dragging and adjusting visual elements, such as axes on parallel coordinate plots. This "what-you-see-is-what-you-get" interactive approach enables domain knowledge to be directly injected into the feature engineering process, facilitating the rapid validation of hypotheses regarding key discriminative features.

### ***2.2.2 Visual Analytics-Guided Iterative Machine Learning***

This technique deeply visualizes the training and evaluation processes of machine learning models and allows users to guide these processes through interactive feedback. For example, in a classification task, the system can visually display decision boundaries, sample confidence levels, and misclassified samples. Users can then interactively re-label uncertain samples, adjust loss function parameters, or add/remove training data, after which the model updates online and visualizes the new results. This human-machine collaborative iterative process aims to optimize model performance while simultaneously enhancing the user's understanding of both model behavior and data characteristics.

### ***2.2.3 Progressive Pattern Construction and Hypothesis Verification Workflow***

Pattern discovery is regarded as a progressive knowledge construction activity. The system provides a series of interactive tools spanning from data querying and pattern sketching to relationship confirmation. Starting from an initial observation of a visual pattern (such as a trend line or a cluster), users can progressively refine it by interactively adding constraints, filtering outliers, or associating other variables, while observing its robustness in real time. The system simultaneously supports the recording of different exploration paths and their corresponding hypotheses, facilitating users in comparing and evaluating the plausibility of various explanations. This thereby transforms exploratory data analysis into a structured hypothesis verification workflow<sup>[3]</sup>.

## ***2.3 Uncertainty Quantification and Visual Reasoning of the Decision-Making Process***

### ***2.3.1 Visual Encoding and Fusion Methods for Sources of Uncertainty***

Uncertainty may arise from data noise, measurement error, model confidence, or sampling

variability. Visual encoding must seamlessly integrate these multi-dimensional uncertainty sources into the primary visual representation. Commonly used methods include employing fuzzy boundaries, gradients of color or luminance, variations in point size or jitter, and confidence interval bands. More advanced techniques utilize aggregated visualization approaches, such as probability volume rendering or the ensemble display of multiple possible outcomes (e.g., hypothesis outcome sets), to convey the overall impact of uncertainty on the prediction of future states.

### ***2.3.2 Visual Traceability and Comparative Analysis of Decision Paths***

To support decision review and auditing, the system must record users' key steps, parameter choices, and intermediate conclusions throughout the interactive analysis process, forming traceable decision paths. These paths can be visualized in forms such as timelines, tree diagrams, or flowcharts. Users can not only retrace their own analytical journey but also concurrently load and compare decision paths from different analysts or under different assumptions. This allows for the intuitive identification of key divergence points that lead to differing conclusions, thereby deepening the understanding of the complexity inherent in the decision problem.

### ***2.3.3 Visual Interactive Interface for Scenario Simulation and Reasoning Analysis***

Scenario simulation oriented toward "what-if" analysis represents an advanced function of decision support. The visual interface allows users to interactively adjust key decision variables, constraints, or external environmental parameters (using sliders, input boxes combined with visual feedback). Based on embedded computational models, the system performs real-time reasoning and visualizes the outcome distributions, changes in key performance indicators, and risk profiles under different scenarios. This reasoning analysis transforms decision-making from a judgment of a single static outcome into an exploration of multiple possible futures and their associated trade-offs, aiding in the identification of robust strategies and the formulation of contingency plans.

## **3. Construction and Evaluation of Interactive Decision Support Systems**

### ***3.1 Decision-Maker-Centered System Integration Design Principles***

#### ***3.1.1 Adaptive Workflow and Modular Service Encapsulation***

The system should provide a dynamically composable analytical workflow engine, encapsulating data processing, feature computation, model inference, and visual rendering as loosely-coupled microservices. Based on the current task phase, decision-makers should be able to flexibly invoke and connect different service modules through a visual orchestration interface to form customized analytical pipelines. The system needs to support the saving, reuse, and sharing of workflows, and be capable of recording historical workflow sequences as repeatable analysis protocols, ensuring the transparency and reproducibility of the analytical process<sup>[4]</sup>.

#### ***3.1.2 Cognitive Load Management and Layered Information Presentation***

In response to the potential information overload from multimodal data, the system needs to integrate proactive cognitive load management mechanisms. This involves the system actively determining the granularity and density of visual information presentation based on the current focus and context of the task. Through strategies of layered information presentation and progressive disclosure, the system initially presents aggregated and abstracted summary views. Only upon receiving users' interactive commands to delve deeper does it progressively provide more detailed raw data, model parameters, or intermediate calculation results. This ensures users' attentional resources remain focused on the information layer most relevant to the current sub-goal of the decision-making process.

#### ***3.1.3 Semantic Unification and Intent Understanding in Multi-Channel Interaction***

The system needs to support and unify natural interaction methods from different channels, such as direct manipulation, gestures, voice commands, and even eye-tracking. The key lies in establishing a unified interaction semantic layer that maps low-level physical interaction events to high-level analytical intents, for example, "compare A with B," "focus on the anomaly region," or "trace the basis for this conclusion." By combining simple interaction context with the task model, the system can perform preliminary inference of user intent, thereby providing more precise visual feedback or automated assistance. This reduces redundant operations and enhances the fluidity of human-machine collaboration.

### ***3.2 Enhancement of Interactive Intelligence in the Decision Support Process***

#### ***3.2.1 Context-Aware Analytical Task Recommendation and Guidance***

By analyzing the current data view, user interaction history, and pending analytical tasks in real-time, the system proactively infers potential subsequent analytical steps in which the user might be interested. For example, when a user performs frequent brushing operations within a cluster boundary region, the system can recommend conducting a difference test or an associated feature analysis for that region. Such recommendations are presented in a non-intrusive manner, such as through subtle visual cues or optional shortcut operations. The algorithm design for this functionality needs to balance the timeliness and accuracy of the recommendations to avoid interrupting the user's dominant train of thought.

#### ***3.2.2 Online Collaborative Model Optimization Based on Interactive Feedback***

The system allows users to utilize judgments formed during the interaction process—such as the reclassification of data points or corrections to model output results—as real-time feedback for online adjustment and optimization of the underlying analytical models. This can manifest as an interactive active learning process, where users label samples about which the model is uncertain to improve its performance. Alternatively, it can enable users to manually adjust model decision thresholds or rule weights and immediately observe the impact on the overall result distribution. This process directly encodes users' domain knowledge into the model, facilitating the co-evolution of the computational model and the user's mental model.

#### ***3.2.3 Explainable Intelligent Assistance and Automated Insight Generation***

To ensure decision-makers' trust in system assistance, any automatically generated suggestions or insights must be accompanied by an explainable rationale. This requires the system to not only output results but also to visualize the key data features, model decision paths, or comparisons with historically similar cases that support those results. For example, when the system detects a potential anomaly pattern, it should simultaneously highlight the primary data dimensions constituting that pattern and their metrics of deviation from the norm. This enables users to rapidly evaluate the plausibility of the insight rather than accepting a "black-box" conclusion.

### ***3.3 Framework for Evaluating Interactive Decision-Making Effectiveness***

#### ***3.3.1 Multi-Dimensional Metrics for Task Performance and Decision Quality***

The evaluation framework must include measurements of objective task performance, such as decision accuracy, completion time, the number of hypotheses examined, and objective utility metrics of the final solution. Simultaneously, it needs to focus on the quality of the decision-making process, such as the number of alternative solutions considered, the comprehensiveness of information utilized, and the adequacy of uncertainty assessment. This requires the design of controlled experimental tasks or simulation-based evaluation scenarios to quantify the performance differences before and after system intervention.

#### ***3.3.2 Evaluation of Cognitive Processes and Subjective User Experience***

Effectiveness evaluation must delve into the decision-maker's cognitive processes. Methods such as retrospective protocol analysis, eye-tracking, or cognitive load scales should be employed to assess the system's role in reducing cognitive load, facilitating pattern recognition, and supporting the construction of reasoning chains. Simultaneously, subjective feedback from users regarding the system's controllability, understandability, trustworthiness, and overall usability should be collected to measure the degree of alignment between the system and the user's cognitive model.

#### ***3.3.3 Longitudinal Evaluation of Insight Generation and Knowledge Construction***

The highest-level objective of interactive decision support is to facilitate the generation of novel insights and the construction of structured knowledge. The evaluation framework should incorporate methods for identifying and measuring "insights." This can be achieved, for instance, by analyzing characteristics of "aha moments" within user interaction logs (such as prolonged fixation followed by a series of exploratory operations), or by comparing users' concept maps of the research problem before and after system use. Longitudinal evaluation may involve a retrospective analysis, after a certain time span, of the decision-making processes supported by the system—assessing the logical rigor and the enduring value of the conclusions derived. This provides a comprehensive assessment of the system's

fundamental enhancement of complex problem-solving capabilities.

## **Conclusion**

This study has conducted a systematic review and investigation into interactive decision-making techniques within multimodal big data visual analytics. It has clarified that unified representation of multimodal data and cognitive enhancement serve as the theoretical foundation, constructed a technological architecture encompassing multi-scale visualization, interactive exploration, and uncertainty reasoning, and proposed human-centered system design and evaluation methodologies. The research findings provide clear direction and a basis for developing intelligent analytical tools capable of effectively addressing data complexity and empowering decision-makers in exploration and judgment. However, numerous avenues for deeper exploration remain in this field. Future research should focus on developing more robust and adaptive cross-modal dynamic fusion mechanisms, designing deeper-level human-machine intelligent collaboration paradigms to achieve precise closed loops of intent and feedback, and establishing more universal and rigorous evaluation standards and benchmark testing environments for interactive decision-making effectiveness. This will propel interactive decision support technology toward a future characterized by greater autonomy, enhanced trustworthiness, and deeper insight.

## **References**

- [1] Zhou Yating, Zhang Rong, and Mou Hui. "Research on Data Visualization Design Framework Based on Multimodal Discourse Analysis." *Packaging Engineering* 46.18 (2025): 75-83.
- [2] Zhang Xiaowen, et al. "Interactive Financial Data Decision Support Visual Analytics Driven by DuPont Analysis and Large Language Models." *Journal of Computer-Aided Design & Computer Graphics* 1-14.
- [3] Huang Ran. *Research on Interactive Hierarchical Quantitative Data Visualization Design for Decision-Making Tasks*. 2023. Southeast University, MA thesis.
- [4] Zhang Rupeng. *Integration Technology for Client-Side Multimodal Spatio-Temporal Data Visual Analytics Systems*. 2020. Southwest Jiaotong University, MA thesis.