Transformation of Composition and Gaze Interaction in Noli Me Tangere Depictions from 1300–1600

Pepe Ballesteros Zapata[1], Nina Arnold[2], Vappu Lukander[3], Ludovica Schaerf[1] and Dario Negueruela del Castillo[1]

[1]*Max Plank Society - University of Zurich, Switzerland*
[2]*University of Zurich, Switzerland*
[3]*University of Helsinki, Finland*

Abstract
This paper examines the development of figure composition and gaze dynamics between Mary Magdalene and Christ in Italian *noli me tangere* depictions from 1300 to 1600 in the context of the emergence of perspective painting. It combines a conceptual, interpretative approach concerning the tactility of the gaze with a compositional analysis. This preliminary study analyzes 51 iconographical images to understand how the gazes between Mary and Christ evolve from pre-perspective to perspective artworks. We estimate gaze direction solely from landmark points, following the assumption that the gaze direction can be estimated from the overall face orientation. Additionally, we develop a metric to quantify the degree of visual interaction between the two protagonists. Our results indicate that Christ is consistently depicted gazing down towards Mary, while Mary displays a broader range of gaze directions. Before the introduction of perspective, the gaze of figures was often rendered solely through face orientation. However, with the advent of the high renaissance, artists began to use complex gestures that separated head orientation from the line of sight.

Keywords
digital art history, gaze estimation, italian painting, noli me tangere

1. Introduction

This study builds upon an analysis of Fra Angelico's *noli me tangere* fresco, dated 1438–1443, which highlights the intrinsic tactility of the gaze between Mary Magdalene and Christ. *Noli me tangere* depictions illustrate the resurrected Christ appearing to Mary on Easter Sunday. Mary reaches out to Christ, who looks at her but gestures for her not to touch him, saying "noli me tangere" (do not touch me) [4].[1] While the iconography translates Christ's words into a visual gesture, the gaze plays a key role. Early Church Fathers already saw an analogy between touch and gaze, interpreting Mary's gaze as recognition [2]. Art historian Barbara Baert then suggests that this exchange of looks replaces physical interaction with a spiritual one, triggering an introspective vision that conveys Mary's understanding of the resurrection [2].

CHR 2024: Fifth Conference on Computational Humanities Research, December 4–6, 2024, Aarhus University, Denmark
jose.ballesteroszapata@uzh.ch (P. B. Zapata); nina.arnold2@uzh.ch (N. Arnold); vappu.lukander@helsinki.fi (V. Lukander); ludovica.schaerf@uzh.ch (L. Schaerf); dario.negueruela@uzh.ch (D. N. d. Castillo)
0000-0002-0720-3535 (P. B. Zapata)

[1]Joh. 20:17.

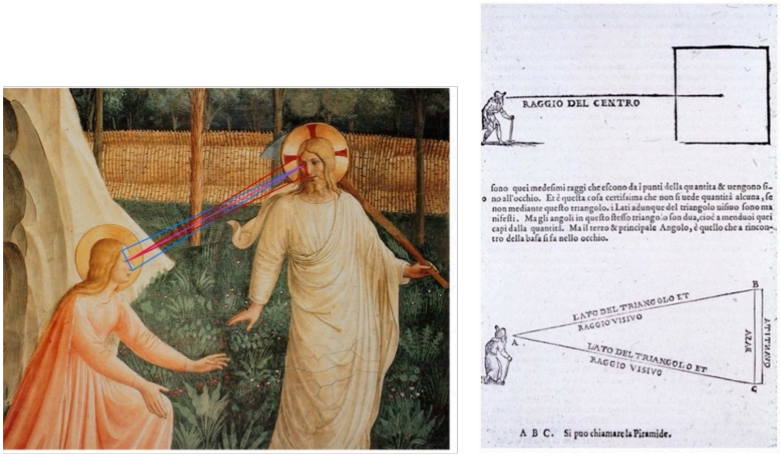

Figure 1: Left: Detail with the visual pyramid and rays of Christ (blue) and Mary (red), and the convergent central ray (pink). Fra Angelico, *noli me tangere*, 1438–1443, Florence, Museo di San Marco. Right: Central ray and visual pyramid, in: L.B. Alberti, *Opuscoli morali*, Venice.

Fra Angelico created his fresco in the early 15th-century Florentine Renaissance, a period marked by the rise of perspective painting [10, 7, 1]. Leon Battista Alberti further demonstrated architect Filippo Brunelleschi's experiment with realistic perspective presentation in "De Pictura", explaining perspective construction through the theory of visual rays [1, 6]. These rays capture sight as lines between the eye and the surface, forming a visual pyramid. The central ray, representing the most intense point of view, is closely connected to the focused object [1].[2] Aligning with the visual center, these lines form the basis of perspective construction [1, 3].[3] Applying this theory to Fra Angelico's fresco, Mary and Christ are understood as actively seeing subjects, their gazes forming visual pyramids centered on each other as the central rays between their eyes align congruently (Fig. 1). This alignment, in the context of 15th-century optical laws, could be interpreted as a quantifiable, albeit invisible, presence. The hypothesis is that with the advent of perspective, Mary and Christ's gazes show greater alignment.

Despite extensive research on the image type within Christian iconography, the scene has not been examined through the lens of perspective and its effects on the composition between Mary and Christ, representing a new area of research. A closely related work found in the literature is the paper by Madhu et al. on human pose and gaze estimation, where Max Imdahl's compositional theories are operationalized using computer vision to detect action regions, lines, and segment the foreground and background in paintings [9]. This work computes rough estimations using body pose orientation to detect the central focus within the image. On the contrary, the proposed method computes single direction estimations from face orientations, and proposes a metric to quantify the degree of visual interaction between two characters present in a painting.

[2]The pyramid consists of outer and middle visual rays and the central ray. The outer rays extend from the observer to the edge of the focused object. The middle rays strike the surface and the central ray forms the core of the pyramid, which strikes the object at a right angle. See Alberti & Bätschmann 2000, p. 203–209.

[3]Alberti's explanations of the visual pyramid are based on Euclid's knowledge of optics, which he applied to scientific painting. See Alberti & Bätschmann 2000, p. 62–65, 72.

By examining the visual rays between Mary and Christ across a selection of iconographical works, this paper enhances the art historical understanding of the compositional interaction between Mary and Christ. Our dataset includes *noli me tangere* images from 1300 to 1600, encompassing both pre-perspective images and the establishment of perspective in painting. This comparative approach aims to understand changes in composition over time.

2. Methodology

2.1. Gathering a dataset

Our dataset is compiled through a combination of filtered data extraction from WikiArt and OmniArt, supplemented by manual research. To avoid imprecision in our study and maintain a focused line of research, only Italian artworks are considered. Hence, we filtered both datasets by school and period of time, selecting Italian paintings from 1300 to 1600. To find our target images, we filtered the resulting images by title, using keywords such as "noli", "resurrection", and "Christ and Mary", finally sorting a total of 39 *noli me tangere* images. To enhance and balance our data, we conducted a manual search in other cultural heritage databases (e.g., Easydb, Bildindex Foto Marburg, and Prometheus). This effort results in a collection of 51 iconographical images: 17 from the 14th century, 15 from the 15th century, and 19 from the 16th century. As we gather data from multiple sources, we ensure the consistency of the metadata by homogenizing it across all the works. As the project explores the composition and gaze evolution of *noli me tangere*, we only take into account the date information of all gathered artworks for our analysis. This dataset, though limited in scope and not intended as historically representative, offers a balanced selection for iconographical study and helps us understand the compositional conventions present in this iconography.

2.2. Extracting facial features

To investigate the gaze direction of iconographical figures, we begin by employing facial landmarks as geometric descriptors to determine the orientation of the face. We use Mediapipe model [8] to extract facial landmarks. Facial recognition in paintings remains a challenging task, as such models were trained on natural images. The digital art history community lacks models fully trained on annotated paintings. To tackle this data discrepancy, we work on cropped images of the faces as the input to the landmark detection model. This approach allows us to bypass the performance limitations of the face recognition model. To extract the bounding box of each face, we use the VGG Image Annotator tool.[4] After the cropped face is processed by Mediapipe's landmark detection model, we collect the following landmark points: the tip of the nose, the chin, the point between the eyes and the outer left-right eye points. To convert the relative image coordinates of facial landmark points to the original painting's image coordinates, we add the top-left corner of each face's bounding box to the corresponding facial landmark coordinates. For each painting, we perform the feature extraction for our two main characters: Mary Magdalene and Christ. Using Mediapipe landmark extraction model, only 17 paintings were successfully processed.

[4]See VGG Image Annotator, 2024: https://www.robots.ox.ac.uk/vgg/software/via/via_demo.html

To overcome this limitation and further process the rest of the paintings in our dataset, we proceed with a manual annotation strategy. We utilize makesense.ai tool to annotate the above listed facial features ourselves.[5] The mid-eye point was not annotated but calculated as the midpoint of the euclidean distance between the eyes. This approach is preferred because annotating the outer eye landmarks, which provide clear reference points, is more reliable than annotating the mid-eye point, which is more challenging to pinpoint. Accurate mid-eye point estimation is crucial as it directly influences the calculation of the central face axis and, consequently, the gaze direction. When the face is oriented completely sideways, the visible eye is used as the mid-eye point.

2.3. Gaze direction estimation

The extracted landmark points give us information about the orientation of the face in the 2D plane. The proposed gaze direction estimation method assumes that the gaze direction corresponds to the overall orientation of the face. The actual gaze direction can be influenced by factors not captured by this method, such as eye movement independent of head orientation. While gaze direction estimation is a well-known task in computer vision, mainly focused on human-computer interactive tools, its implementation requires large amounts of training data not suitable for the nature of the data at hand [5]. We address this issue by estimating the gaze direction based solely on landmark points. The estimated gaze vector reflects Alberti's concept of the central ray, as the direction which represents the strongest line of sight to the focused object. To estimate the 2D gaze direction for each face, we start by forming a central face axis $\vec{c}_{\text{face}}$ using the mid-eye and chin landmark points:

$$\vec{c}_{\text{face}} = \begin{pmatrix} x_{\text{chin}} - x_{\text{mid-eye}} \\ y_{\text{chin}} - y_{\text{mid-eye}} \end{pmatrix} \tag{1}$$

This axis is sensitive to pitch and roll face rotations, contained in the 2D image plane. As we want to calculate directions in 2D, we do not need to take into account the jaw rotation of the face, as it only affects the gaze direction along the axis outside the 2D plane. To estimate the gaze direction from the central axis of the face, we use the tip of the nose as a reference point for face orientation. The gaze direction $\vec{g}$ is defined as the vector perpendicular to the central facial axis that intersects the tip of the nose. This vector can be found rotating $\vec{c}_{\text{face}}$ by 90 degrees and performing a translation to the tip of the nose.

$$\vec{g} = (x_{\text{nose}}, y_{\text{nose}}) + \begin{pmatrix} -(y_{\text{chin}} - y_{\text{mid-eye}}) \\ x_{\text{chin}} - x_{\text{mid-eye}} \end{pmatrix} \tag{2}$$

In a scenario where our two protagonists look at each other, the gaze direction vectors are parallel. We create a vector, called the baseline vector $\vec{b}$, formed by the line that passes through the tip of their noses:

$$\vec{b} = \begin{pmatrix} x_{\text{nose2}} - x_{\text{nose1}} \\ y_{\text{nose2}} - y_{\text{nose1}} \end{pmatrix} \tag{3}$$

[5]See makesense.ai, 2024, https://www.makesense.ai.

We use the baseline vector $\vec{b}$ as a reference of direction that represents the line of sight when the two characters look at each other. If the gaze direction vectors are aligned with the baseline vector, it confirms that both figures are looking directly at each other, establishing a visual connection. To quantify the angular deviation of each protagonist's gaze from this reference line we use the cosine similarity. The cosine similarity is calculated as the dot product of the normalized gaze direction $\vec{g}$ and baseline $\vec{b}$ vectors :

$$\cos\theta = \frac{\vec{g}\cdot\vec{b}}{\|\vec{g}\|\|\vec{b}\|} \tag{4}$$

This metric provides a numerical representation for the degree of visual connection between characters, producing values between -1 (looking away from each other) and 1 (direct mutual gaze). By calculating two cosine similarity values, one for each character, we can not only assess mutual connection but also determine to what extent each character is looking at the other. A high cosine similarity value for Christ's gaze direction indicates that Christ is looking towards Mary, and conversely, a high value for Mary's gaze direction signifies that she is looking towards Christ.

3. Results

Our hypothesis suggests that the introduction of perspective in Italian painting leads to a greater alignment in the gazes of Mary and Christ, quantified using the cosine similarity between a baseline vector and the gaze direction of each character. Contrary to our expectations, our findings do not support a temporal trend consistent with the hypothesis; the cosine similarity rates do not increase as perspective becomes more established in Italian painting. As illustrated in Fig. 2 the visual connection between Mary and Christ becomes less consistent starting from the advent of the high renaissance in the late 15th century onwards. This divergence is largely due to the strong assumption of our methodology, which is challenged by the later high renaissance depictions, where intricate gestures and expressions complicate the alignment of gaze with body orientation. Consequently, our approach is not sensitive to the nuances of perspective in later Italian renaissance artworks. However, in analyzing the gaze dynamics between Mary and Christ within our dataset, we observe distinct patterns. On average, Mary exhibits a mean cosine similarity of 0.80, while Christ demonstrates a higher mean of 0.91. This indicates that Christ is more frequently depicted gazing towards Mary, whereas Mary is more likely looking elsewhere. Notably, we identified three paintings where Mary's cosine similarity falls below 0.4, where she looks at Christ's feet in an adoration pose. In such instances, Christ consistently gazes towards her with a cosine similarity above 0.75. The high cosine similarity values for Christ across the three centuries underscore a thematic consistency within the iconography of Christ looking at Mary. The appearance of low cosine similarity values throughout our dataset can be attributed to more rudimentary pre-perspective depictions (Fig. 3 on the left) but also to the strong model assumptions (Fig. 3 on the right). A visual and qualitative evaluation suggests that gazes that clearly deviate from the baseline direction (cosine similarity less than 0.7) are mainly from the 15th and 16th centuries (14th century: 2; 15th

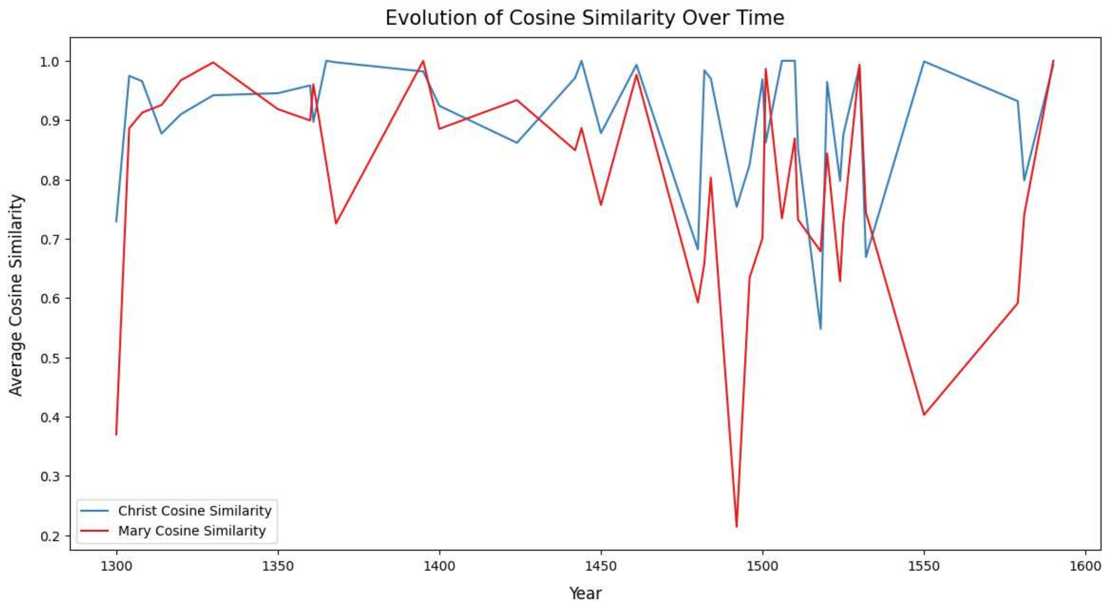

Figure 2: Evolution of cosine similarity for both protagonists in *noli me tangere*. The cosine similarity measures the degree of visual connection for each protagonist.

century: 6; 16th century: 6). This observation suggests that the proposed method face challenges in accurately discerning the gaze direction in images with a higher degree of structural complexity. The image on the right in Fig. 3 provides an example of the model's limitations. The main contributing factor for this limitation is the independent movement of the eyeballs with respect to the face orientation and body gesture. Complex head gestures, particularly found in the advent of the high renaissance, are more likely to produce a discrepancy between facial orientation and gaze direction. However, Fig. 4 shows examples of both pre-perspective and perspective depictions where the model has calculated the gaze directions reasonably well. The gestures of the figures are relatively simple, and the gaze direction correlates with the orientation of the face, thus supporting our previously mentioned findings on the limitations of the model. These illustrative examples show that the model performs better when the image features a lower level of structural complexity.

A significant compositional convention observed is Christ's tendency to gaze downwards towards Mary. Conversely, Mary primarily looks downwards only when looking at Christ's feet, while her gaze directions exhibit greater variability. Specifically, our model indicates that Christ looks downwards 50 times and upwards once, whereas Mary looks upwards 37 times and downwards 14 times (Fig. 5).

We calculate the position of Mary and Christ in each painting by using the x-coordinate of the outer right eye point of each figure, which allows us to categorize their placement as either left or right. The analysis suggests a moderate compositional convention regarding figure placement. Across our dataset, Christ is positioned to the right 36 times and to the left 15 times. Interestingly, this convention weakens over time. In late medieval depictions Christ predominantly appears on the right, with only one instance on the left out of 17 analyzed

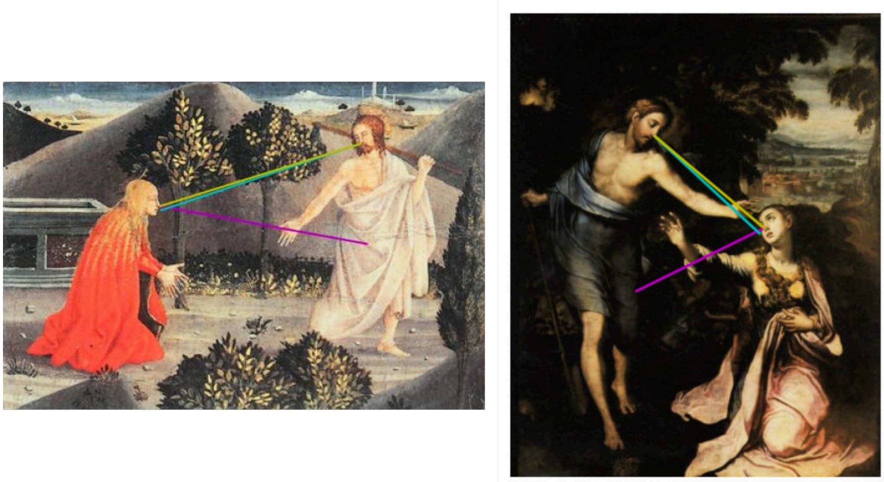

Figure 3: Visualization of gaze direction vectors calculated with the proposed method. Each vector is represented for the tip of the nose for Mary Magdalene (purple) and Christ (blue). The baseline vector is represented in yellow. Left: example of gaze estimation limitations due to rudimentary pre-perspective depictions. Piero della Francesca, *noli me tangere*, 1444, Misericordia Altarpiece predella, Sansepolcro, Pinacoteca Comunale. Right: another limiation example due to our model assumption in a later depiction. Notice the difference in gesture complexity between the two images. Marco del Pino, *noli me tangere*, c. 1550, Museo Cappuccini, Rome.

Figure 4: Examples of high cosine similarity values indicating gaze alignment. Left: example from an early depiction. Mariotto di Nardo, *Scenes From the Life of Christ, noli me tangere*, 1395, Musée du Petit Palais, Avignon. Right: example of a later depiction. Jacopo Pontormo, *noli me tangere*, 1530, Private collection.

images. This tendency becomes less pronounced in the 15th century, where 5 out of 15 images depict Christ on the left. By the 16th century, 9 out of 19 paintings depict Christ on the left, suggesting that artists exercised greater freedom in determining figure placement.

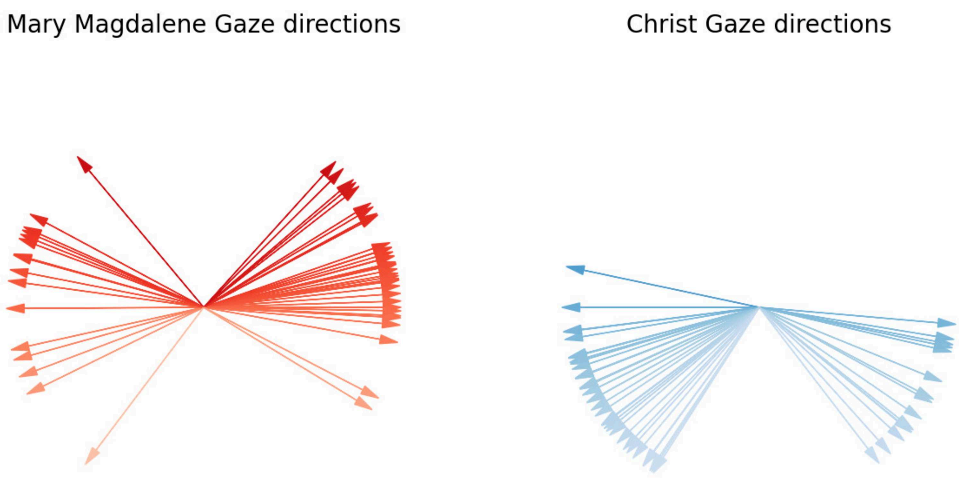

Figure 5: A comparison of gaze directions between Mary Magdalene and Christ in *noli me tangere* depictions. The gaze direction of a single painting is represented as a vector of unit length for each character.

4. Conclusion

This preliminary work presents a computational method to tackle the quantification of gaze interactions in paintings. Particularly, this work is focused on Italian *noli me tangere* depictions in the context of the emergence of perspective. The estimated gaze direction vector, echoing Alberti's central ray of sight, is calculated using solely facial landmark points under the assumption that the line of sight can be well represented as the overall face orientation direction. However, results reveal that face orientation alone is insufficient to capture the gaze direction in artworks where figures are depicted in complex gestures. This is particularly evident in most images from the 15th century onwards. As perspective developed, so did the complexity of paintings, as reflected in the postures of the figures and compositional conventions. Results show that before the introduction of perspective, the gaze of figures was often depicted through face orientation. The introduction of perspective allowed artists to convey the feeling of looking while experimenting with more complex gestures, where the line of sight became more independent from the face and body's gestures. The depiction of gaze became possible through more nuanced elements, such as subtle misalignments between body, face, and line of sight.

While our proposed method remains insensitive to perspective, it allows the exploration of compositional conventions within the Italian *noli me tangere* depictions, revealing patterns in the placement and gaze dynamics of Christ and Mary. Specifically, we found that Christ is consistently depicted gazing towards Mary, while Mary displays a more spread range of gaze directions. Christ's steadfast gaze on Mary across the centuries reflects his central role as the sole active communicator in the scene. His focused gaze symbolizes the divine insight conveyed to Mary about his resurrection and delivers a spiritual message without words. In the depictions of *noli me tangere*, the theological discussion about the actual physicality of the resurrected Christ is overshadowed by the emphasis on the visionary gaze and the act of not touching. Here, there is no need for touch to believe. In contrast, Mary's more variable gaze directions mirror her human state of confusion about identifying the Resurrected, whom she recognizes as Christ only after he turns to her and says "noli me tangere." Her gazes also reflect emotional poignancy, ranging from intense staring to a worshipful pose with downcast eyes. Mary embodies the complexity of the Christian faith, which is based not on her knowledge, but on her openness to visionary revelation and her trust. Moreover, our analysis indicates a shift in compositional conventions over time, with artists exercising greater freedom in figure placement, further contributing to the complexity present in this iconography.

Future work focuses on relaxing the assumptions of the proposed methodology. For example, using high-resolution models to enhance the resolution of cropped faces would facilitate the manual annotation of pupils. This improvement will enable a clearer disentanglement between face orientation and gaze direction. Moreover, extracting 3D Morphable Models (3DMM) from the 2D facial keypoints could allow the reconstruction of the 3D geometry of the scene, offering a comprehensive understanding of the spatial relationships between characters. The measuring of distances, angles and orientations in a 3D space can provide sensitive information to depictions that use perspective. Further expansions of the iconographical dataset might incorporate the German and Flemish schools to broaden the scope and comparative dimensions. Additionally, integrating older *noli me tangere* images, dating back to the Carolingian era and executed in a distinctly two-dimensional style like manuscript illuminations, could elucidate the developmental process and identify further stylistic evolutions. Other future lines include further exploration of compositional aspects, such as analyzing the gesture of Christ's repelling motion and Mary's yearning pose. While this research has focused on paintings, it would be intriguing to examine these themes in reliefs. Due to their material nature, reliefs present unique challenges in perspective that could enrich our understanding of artistic expression across different media.

References

[1] L. B. Alberti and O. Bätschmann. *Das Standbild; Die Malkunst; Grundlagen der Malerei.* Darmstadt: Wissenschaftliche Buchgesellschaft, [1435/36] 2000.

[2] B. Baert. *To touch with the gaze: Noli me tangere and the iconic space.* Leuven, 2012.

[3] H. Belting. *Florenz und Bagdad: eine west√∂stliche Geschichte des Blicks.* 3. Aufl., [Neuausg. 2012]. Beck'sche Reihe 6019. München: C.H. Beck, 2012.

[4] A. Beriger, W.-W. Ehlers, and M. Fieger. *Hieronymus, Biblia Sacra Vulgata: Lateinisch-deutsch.* Sammlung Tusculum. Berlin: De Gruyter, 2018.

[5] D. Cazzato, M. Leo, C. Distante, and H. Voos. "*When i look into your eyes: A survey on computer vision contributions for human gaze estimation and tracking*". In: *Sensors* 20.13 (2020), p. 3739.

[6] S. Y. Edgerton. *The Renaissance rediscovery of linear perspective.* Icon editions. New York: Evanston, 1976.

[7] A. González Mozo. “”Strategies for Depicting Sacred Stories””. In: *Fra Angelico and the rise of the Florentine Renaissance.* Madrid: Museo Nacional del Prado, 2019, pp. 173–195.

[8] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C.-L. Chang, M. G. Yong, J. Lee, W.-T. Chang, W. Hua, M. Georg, and M. Grundmann. *MediaPipe: A Framework for Building Perception Pipelines.* 2019. arXiv: 1906.08172 [cs.DC]. URL: https://arxiv.org/abs/1906.08172.

[9] P. Madhu, T. Marquart, R. Kosti, P. Bell, A. Maier, and V. Christlein. “*Understanding Compositional Structures in Art Historical Images using Pose and Gaze Priors*”. In: *arXiv.org* (2020).

[10] C. B. Strehlke and A. González Mozo. *Fra Angelico and the rise of the Florentine Renaissance.* Madrid: Museo Nacional del Prado, 2019.