LCKGE: Multimodal Knowledge Graph Embeddings via Lorentz-based Contrastive Learning

Appendix A Related Works

The study of knowledge graph embedding has far-reaching implications for the fields of knowledge representation and applications. Existing KGE methods are divided into two main branches: unimodal KGE methods and multimodal KGE methods, from the perspective of the number of modalities involved in knowledge graph embedding.

A.1 Unimodal Knowledge Graph Embedding Models

Translation-based KGE methods are popularized by TransE [1], which interprets the entities and relations as points and translation transformations in low-dimensional vector space, *e.g.* $\boldsymbol{h} + \boldsymbol{r} \approx \boldsymbol{t}$. Subsequently, to solve the one-to-many problem, TransH [2] proposes the relation-specific hyperplanes with normal vector. And TransR [3] points out that the intrinsic distribution of entities and relations are different, so it projects them into separate spaces. There are other extended variants along this line, *e.g.* TransD [4], which uses independent projection vectors to reduce the amount of parameters. RotatE [5] first extends the embedding space to complex space, and models the relation mapping as 2D rotation transformations. Rotate3D [6] models the relations as 3D rotations. For further exploration, OTE [7] generalizes the n-dimensional rotation as an orthogonal matrix. Besides the translation-based method, many other effective approaches have been explored. Semantic matching approaches include bilinear models such as RESCAL [8], which uses bilinear operations to model semantic interactions, DistMult [9], which restricts the parameters of the relational matrix to the diagonal, and ComplEx [10], which extends the scoring function to the complex space and QuatE [11], which proposes to use the quaternion numbers for more complex interactions.

Embedding KGs in manifolds is another line. ManifoldE [12] first introduces the concept of manifolds in KGE extending pointwise embeddings to manifold-based embeddings, and MöbiusE [13] embeds KGs on Möbius ring.Recently, several works [14, 15, 16] have proved that *Riemannian* geometric manifolds are more expressive than Euclidean spaces for the task of embedding graph-structured data. Therefore, MuRP [17] first embeds KGs in hyperbolic space, and AttH [18] eliminates manual curvature setting to further enhance embedding capabilities. Further, GIE [19] and M^2GNN [20] introduce mixed-curvature manifolds for modeling more complex structures.

A.2 Multimodal Knowledge Graph Embedding Models

Beyond the unimodal KGE models, a number of multimodal KGE methods have been proposed recently, which exploit additional multimodal information of entities to further improve the performance of the link prediction task.

To encode image features in KG embeddings, IKRL [21] learns visual and structural information separately, and uses a TransE model [1] for decoding. While TransAE [22], inspired by the architecture of the auto-encoder decoder, proposes to jointly represent structural and multimodal features in a unified latent space. RSME [23] notices the damage of multimodal information on structure embeddings and proposes a forget gate to filter out the multimodal data with negative impact. MKBE [24] further encodes multimodal information *e.g.* visual, linguistic and date among others. Then OTKGE [25] uses an Optimal Transfer (OT) scheme to fuse multimodal and structural embeddings. MoSE [26] proposes a Tight-Coupling Relation (TCR) decoupling strategy to resolve intermodal contradictions, while IMF [27] proposes a bilinear fusion module to increase the effective interaction between structural and multimodal information. With Transformer [28] attracting more attention, MKGformer [29] introduces Transformer architecture for multimodal fusion. In addition, other recent works focus on multimodal knowledge graphs on recommender systems, *e.g.* MMGCN [30], MMGAT [31].

Appendix B Preliminaries

B.1 Hyperbolic Space

Hyperbolic space is a type of constant negative curvature *Riemannian* manifold with five isometric models, *e.g.* Poincaré sphere, Poincaré half-plane, Klein model, and hyperboloid (Lorentz) model, therefore any point in hyperbolic space can be equivalently transformed to these models using conformal transformations. Considering efficient parameterization and numerical stability, the Lorentz model is used as the cornerstone of our work, so an n-dimensional Lorentz model $\mathbb{L}^n$ is defined as follows

$$\mathbb{L}^n := \{\boldsymbol{x} = [x_0, \cdots, x_n] \in \mathbb{R}^{n+1} | \langle \boldsymbol{x}, \boldsymbol{x} \rangle_{\mathcal{L}} = -1, x_0 > 0\} \tag{B-1}$$

where $\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\mathcal{L}} = \boldsymbol{x}^\top \boldsymbol{\mathcal{G}} \boldsymbol{y}$ and $\boldsymbol{\mathcal{G}} = \mathrm{diag}(-1, \mathbf{I}_n)$, $\mathbf{I}_n$ is $n \times n$ identity matrix and $\mathrm{diag}(\cdot)$ denotes a diagonal matrix.

B.2 Lorentz Transformations

In the n-dimensional Lorentz model $\mathbb{L}^n$, the Lorentz transformations $\Phi : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ is a family of linear transformations with $2n$ degrees of freedom if $\langle \Phi(\boldsymbol{x}), \Phi(\boldsymbol{y}) \rangle_{\mathcal{L}} = \langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\mathcal{L}}$ is hold for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^{n+1}$. All Lorentz transformations Φ form a group $\mathbf{O}(1, n)$ named as *Lorentz Group*, which is defined as

$$\mathbf{O}(1, n) := \{\boldsymbol{\Gamma} \in \mathrm{GL}(1 + n, \mathbb{R}) | \boldsymbol{\Gamma} \boldsymbol{\mathcal{G}} \boldsymbol{\Gamma}^\top = \boldsymbol{\Gamma}^\top \boldsymbol{\mathcal{G}} \boldsymbol{\Gamma} = \boldsymbol{\mathcal{G}}\} \tag{B-2}$$

where $\mathrm{GL}(1 + n, \mathbb{R})$ is the general linear group of $(n + 1) \times (n + 1)$ invertible matrices over real number domain. Thus, any point in Lorentz model can be transformed with matrix multiplications, *i.e.* $\hat{\boldsymbol{x}} = \boldsymbol{\Gamma} \cdot \boldsymbol{x}, \boldsymbol{x} \in \mathbb{L}^n, \boldsymbol{\Gamma} \in \mathbf{O}(1, n)$.

In geometric, any Lorentz transformation $\boldsymbol{\Gamma} \in \mathbf{O}(1, n)$ can be decomposed into a combination of a Lorentz Boost $\mathbf{B}$ and a Lorentz Rotation $\mathbf{R}$ by polar decomposition [32], which is described as following

$$\boldsymbol{\Gamma} = \mathbf{R} \cdot \mathbf{B} = \begin{bmatrix} 1 & \mathbf{0}^\top \\ \mathbf{0} & \hat{\boldsymbol{R}} \end{bmatrix} \begin{bmatrix} \gamma & -\gamma \boldsymbol{v}^\top \\ -\gamma \boldsymbol{v} & \boldsymbol{I}_n + \frac{\gamma^2}{1+\gamma} \boldsymbol{v} \boldsymbol{v}^\top \end{bmatrix} \tag{B-3}$$

where $\hat{\boldsymbol{R}} \in \mathbf{SO}(n)$ is an orthogonal matrix, *i.e.* $\hat{\boldsymbol{R}}^\top \hat{\boldsymbol{R}} = \boldsymbol{I}$ and $\boldsymbol{v} \in \mathbb{R}^n$ is velocity of each axis, $\gamma = \frac{1}{\sqrt{1 - \|\boldsymbol{v}\|^2}}$. The Lorentz rotation $\mathbf{R}$ is the rotation of the spatial coordinates and the Lorentz boost $\mathbf{B}$ describes relative motion with constant velocity. in this work, we utilize these two types of transformations to learn relation mappings.

B.3 Problem Definition

A multimodal knowledge graph (MKG) is defined as $\mathcal{G} = \{\mathcal{E}, \mathcal{R}, \mathcal{T}\}$, where $\mathcal{E}$ is the entity set, $\mathcal{R}$ is the relation set, and $\mathcal{T} = \{(h, r, t) | h, t \in \mathcal{E}, r \in \mathcal{R}\}$ is the set of triplets of the MKG. Here the entity $e \in \mathcal{E}$ is equipped with the *structure* modal embedding $\boldsymbol{e}_s$, which can be learned with symbol triplets, and other given *multimodal* modal embeddings $\boldsymbol{e}_m$, *e.g.* *visual* modal embedding $\boldsymbol{e}_v$, *linguistic* modal embedding $\boldsymbol{e}_l$. Practically, the multi-modal knowledge graphs are usually incomplete, hence completing the link prediction task is a necessary task to develop an effective multimodal KGE method.

B.4 Model Formulation

B.4.1 Definition of Entities and Relations Embedding.

Given a triplet (h, r, t), the entities h, r are vectors in a low-dimensional space while the relation $r \in \mathcal{R}$ is interpreted as two types of Lorentz transformations *i.e.* Lorentz rotation transformation and Lorentz boost transformation $\boldsymbol{\Gamma}_{\mathrm{rot}}, \boldsymbol{\Gamma}_{\mathrm{boost}}$ defined as Eq.B-3. However, the optimization with orthogonal matrix constraint, such as $\hat{\boldsymbol{R}}$ in Eq.B-3, involves manifold learning [33], which is time-consuming. Therefore, for the entity embeddings of each modality, we set corresponding relation-specific Lorentz transformations, *e.g.* $\boldsymbol{\Gamma}^s_{\mathrm{rot}}, \boldsymbol{\Gamma}^s_{\mathrm{boost}}$ for structure modality, $\boldsymbol{\Gamma}^v_{\mathrm{rot}}, \boldsymbol{\Gamma}^v_{\mathrm{boost}}$ for visual modality, and $\boldsymbol{\Gamma}^l_{\mathrm{rot}}, \boldsymbol{\Gamma}^l_{\mathrm{boost}}$ for linguistic modality.

Then the structure modal embeddings of all entities are defined as $\boldsymbol{S} \in \mathbb{R}^{|\mathcal{E}| \times n}$, where $|\cdot|$ is the cardinality of a given set. In addition, the other two additional multimodal embeddings for each entity are the visual modal embeddings $\boldsymbol{V} \in \mathbb{R}^{|\mathcal{E}| \times d_v}$ and the linguistic modal embeddings $\boldsymbol{L} \in \mathbb{R}^{|\mathcal{E}| \times d_l}$, where d_v and d_l are the dimension of the visual and linguistic embeddings, respectively.

Appendix C Benchmarks Statistics

This section mainly introduces three well-established benchmarks used in our experiments, WN9, FB15k-237, and FB15K. Table C-1 shows the statistics of the three benchmarks. The ”Avg. #Degree” column in the table is defined as follows: We first calculate the average degree of each entity under each relation type, and then compute the mean and standard deviation of entity degree across all relation types.

Table C-1: Statistic information of three benchmarks, average node degree plus/minus standard deviation.

Dataset	#Entities	#Relations	#Training	#Validation	#Test	Avg. #Degree	Heterogeneity	Scale
WN9	6,555	9	11,742	1,338	1,320	1.53±0.23	Low	Small
FB15K-237	14,541	237	272,115	17,535	20,466	4.56 ±6.14	High	Medium
FB15K	14,952	1,345	414,550	59,222	118,444	2.40±3.15	High	Large

Appendix D Experiments

D.1 Temperature Factor

The temperature factor τ in Eq.12 is a relatively important parameter for controlling the degree of multimodal fusion. Here we evaluate the temperature factor τ from 0.01 to 3.2 and plot the link prediction performance of LMKGE on the FB15K benchmark with the MRR metric as Figure D-1.

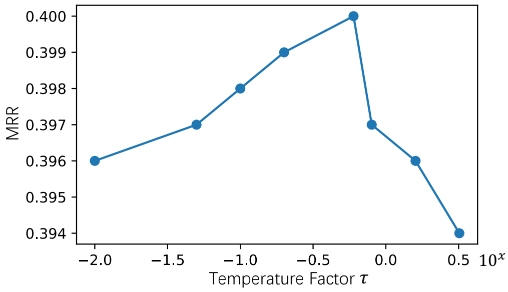


Figure D-1: Effect of different temperature factors τ on the performance of the FB15K-237.

D.2 Embedding Dimension

In this section we mainly discuss the impact of embedding dimensions on model performance on the FB15K-237 benchmark. We set the embedding dimension from 20 to 800, and the MRR metric is used to measure the model performance. The experimental results are shown in Figure D-2.

D.3 Number of the Nearest Neighbors

For multimodal fusion, we propose a nearest neighbor fusion module based on contrastive learning to facilitate intermodal information sharing. We set the intermodal positive samples with size k in Eq.(7) and the multimodal intramodal positive samples with size k' in Eq.(8). Here, we mainly explore the impact of two hyperparameters k and k' on the performance of the model on FB15k-237 with embedding dimension 32. The experimental results are shown in Figure D-3.

D.4 Weight of Multimodal Fusion loss $\mathcal{L}_{\mathrm{CL}}$

To fully explore the impact of the weight of the multimodal fusion module on the performance of the model, we reformulate the optimization loss as $\mathcal{L} = \mathcal{L}_{\mathrm{KGE}} + \alpha \cdot \mathcal{L}_{\mathrm{CL}}$, adjusting the weight of multimodal fusion in the final optimization by introducing a parameter α, and conduct a series of experiments on the FB15K-237 benchmark, varying α from 0 to 2. The experimental results are shown in Figure D-4.

D.5 Analysis of the Complexity

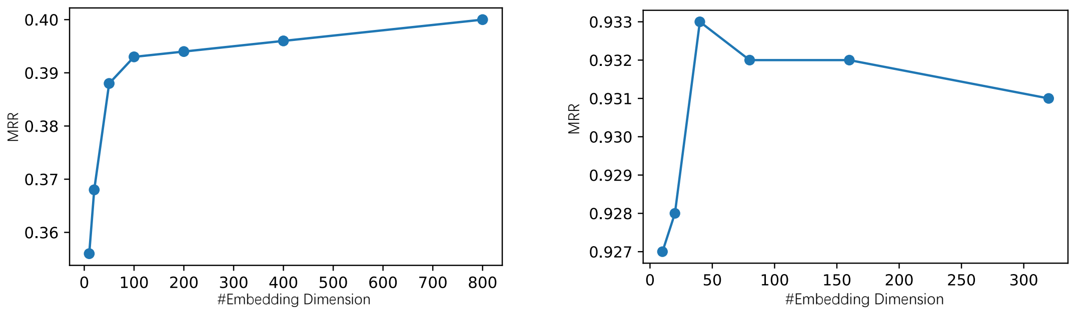
To evaluate the effectiveness of reformulating Lorentz Rotation matrix with quaternion numbers, we perform comparative experiments, comparing the ”Gram-Schmidt” method with the ”Quaternions” method for the orthogonalization of the orthogonal matrix $\hat{R}$ in Eq.(3). The experimental results are shown in Table D-2.

D.6 Embedding Visualizations

To demonstrate the advantages of multimodal representations, we will show the entity embeddings of both unimodal and multimodal methods. Specifically, we will compare the unimodal method, ComplEx, with the LMKGE method that uses the structure modality only, as well as the multimodal method LMKGE method that incorporates multiple modalities on the benchmark FB15K-237, where we visualize the embeddings of entities with three types (”tv”, ”media common”, ”location”), and the results are shown in Figure D-5. The LMKGE model with a single structural modality can produce better entity embeddings compared to ComplEx, while the use of multi-modalities further improves the embedding capabilities for the KGE model to produce superior quality entity representations.

Table D-2: Comparative experiments for ”Gram-Schmidt” and ”Quaternions”

Model	Embedding Dims.	Benchmark	Consuming Time (1 epoch)	#Param	MRR
LCKGE(Gram-Schmidt)	32	FB15k-237	23s	78,573,185	0.382
LCKGE(Quaternions)	32	FB15k-237	17s	78,027,137	0.383
Comparison	-	-	-23.08%	-0.70%	-



(a) Embedding dimension evaluation on FB15k-237. (b) Embedding dimension evaluation on WN9.

Figure D-2: Experimental results evaluating the impact of embedding dimensions on model performance.

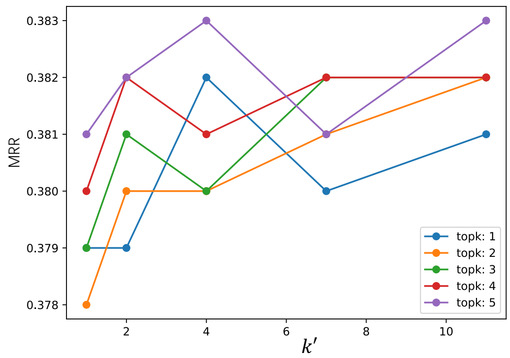


Figure D-3: Effect of different k and k' on the performance of the FB15K-237. The "topk" in the figure is the number of intermodal positive samples k, and the x-axis is the number of intramodal positive samples k'.

Appendix E Method

E.1 Lorentz Rotation

As shown in Eq.(3), the Lorentz transformation can be decomposed into Lorentz rotation and Lorentz boost, where $\hat{R}$ is an orthogonal matrix. Since optimization with orthogonality constraints involves manifold learning, which is time-consuming, we use quaternion numbers to reformulate the orthogonal matrix in Lorentz rotation to reduce computational complexity. Specifically, the Hamiltonian product $\otimes$ for any quaternion $q = a + b\boldsymbol{i} + c\boldsymbol{j} + d\boldsymbol{k}$ and a united $\hat{q} = \hat{a} + \hat{b}\boldsymbol{i} + \hat{c}\boldsymbol{j} + \hat{d}\boldsymbol{k}, \|\hat{q}\| = \sqrt{\hat{a}^2 + \hat{b}^2 + \hat{c}^2 + \hat{d}^2} = 1$ can be reformulated with matrix multiplications as follows

$$\tilde{q} = \hat{q} \otimes q = \begin{bmatrix} \hat{a} & -\hat{b} & -\hat{c} & -\hat{d} \\ \hat{b} & \hat{a} & -\hat{d} & \hat{c} \\ \hat{c} & \hat{d} & \hat{a} & -\hat{b} \\ \hat{d} & -\hat{c} & \hat{b} & \hat{a} \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \tag{E-4}$$

where the left matrix is a 4×4 orthogonal matrix. Therefore, we replace the $\hat{R}$ in Eq.(3) with quaternion number to reformulate the Lorentz Rotation to accelerate inference and relational transformation.

References

[1] Antoine Bordes, Nicolas Usunier, Alberto Garcia-Duran, Jason Weston, and Oksana Yakhnenko, "Translating embeddings for modeling multi-relational data," *Advances in neural information processing systems*, vol. 26, 2013.

[2] Zhen Wang, Jianwen Zhang, Jianlin Feng, and Zheng Chen, "Knowledge graph embedding by translating on hyperplanes," in *Proceedings of the AAAI conference on artificial intelligence*, 2014, vol. 28.

[3] Yankai Lin, Zhiyuan Liu, Maosong Sun, Yang Liu, and Xuan Zhu, "Learning entity and relation embeddings for knowledge graph completion," in *Proceedings of the AAAI conference on artificial intelligence*, 2015, vol. 29.

[4] Guoliang Ji, Shizhu He, Liheng Xu, Kang Liu, and Jun Zhao, "Knowledge graph embedding via dynamic mapping matrix," in *Proceedings of the 53rd annual meeting of the association for computational linguistics and the 7th international joint conference on natural language processing (volume 1: Long papers)*, 2015, pp. 687–696.

[5] Zhiqing Sun, Zhi-Hong Deng, Jian-Yun Nie, and Jian Tang, "Rotate: Knowledge graph embedding by relational rotation in complex space," *arXiv preprint arXiv:1902.10197*, 2019.

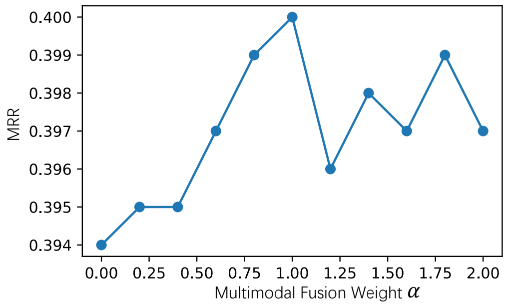
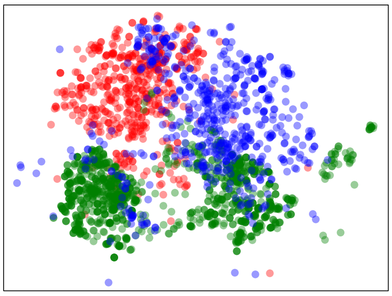
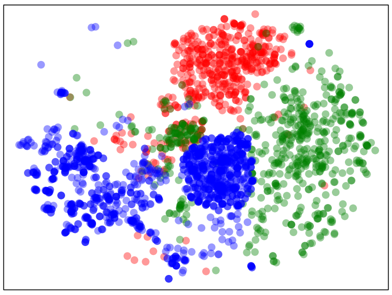


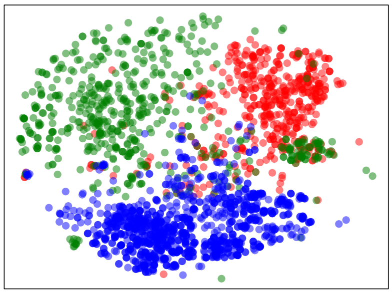
Figure D-4: Experimental results on evaluation of varying the weight α of multimodal fusion loss $\mathcal{L}_{\mathrm{CL}}$ on FB15k-237.



(a) Entities representation of ComplEx



(b) Entities representation of LMKGE with one structure modality



(c) Entities representation of LMKGE with multimodalities

Figure D-5: Visualization of entities embeddings on three KGE methods. Figure 5(a) displays the structure entities embeddings learned by ComplEx, while Figure 5(b) shows the learned entities representations with structure modality in LMKGE. Lastly, Figure 5(c) illustrates the entity representations learned by LMKGE with multimodalities.

[6] Chang Gao, Chengjie Sun, Lili Shan, Lei Lin, and Mingjiang Wang, "Rotate3d: Representing relations as rotations in three-dimensional space for knowledge graph embedding," in *Proceedings of the 29th ACM international conference on information & knowledge management*, 2020, pp. 385–394.

[7] Yun Tang, Jing Huang, Guangtao Wang, Xiaodong He, and Bowen Zhou, "Orthogonal relation transforms with graph context modeling for knowledge graph embedding," *arXiv preprint arXiv:1911.04910*, 2019.

[8] Maximilian Nickel, Volker Tresp, Hans-Peter Kriegel, et al., "A three-way model for collective learning on multi-relational data.," in *Icml*, 2011, vol. 11, pp. 3104482–3104584.

[9] Bishan Yang, Wen-tau Yih, Xiaodong He, Jianfeng Gao, and Li Deng, "Embedding entities and relations for learning and inference in knowledge bases," *arXiv preprint arXiv:1412.6575*, 2014.

[10] Théo Trouillon, Johannes Welbl, Sebastian Riedel, Éric Gaussier, and Guillaume Bouchard, "Complex embeddings for simple link prediction," in *International conference on machine learning*. PMLR, 2016, pp. 2071–2080.

[11] Shuai Zhang, Yi Tay, Lina Yao, and Qi Liu, "Quaternion knowledge graph embeddings," *Advances in neural information processing systems*, vol. 32, 2019.

[12] Han Xiao, Minlie Huang, and Xiaoyan Zhu, "From one point to a manifold: Knowledge graph embedding for precise link prediction," *arXiv preprint arXiv:1512.04792*, 2015.

[13] Yao Chen, Jiangang Liu, Zhe Zhang, Shiping Wen, and Wenjun Xiong, "Möbiuse: Knowledge graph embedding on möbius ring," *Knowledge-Based Systems*, vol. 227, pp. 107181, 2021.

[14] James W Cannon, William J Floyd, Richard Kenyon, Walter R Parry, et al., "Hyperbolic geometry," *Flavors of geometry*, vol. 31, no. 59-115, pp. 2, 1997.

[15] Ines Chami, Zhitao Ying, Christopher Ré, and Jure Leskovec, "Hyperbolic graph convolutional neural networks," *Advances in neural information processing systems*, vol. 32, 2019.

[16] Octavian Ganea, Gary Bécigneul, and Thomas Hofmann, "Hyperbolic neural networks," *Advances in neural information processing systems*, vol. 31, 2018.

[17] Ivana Balazevic, Carl Allen, and Timothy Hospedales, "Multi-relational poincaré graph embeddings," *Advances in Neural Information Processing Systems*, vol. 32, 2019.

[18] Ines Chami, Adva Wolf, Da-Cheng Juan, Frederic Sala, Sujith Ravi, and Christopher Ré, "Low-dimensional hyperbolic knowledge graph embeddings," *arXiv preprint arXiv:2005.00545*, 2020.

[19] Zongsheng Cao, Qianqian Xu, Zhiyong Yang, Xiaochun Cao, and Qingming Huang, "Geometry interaction knowledge graph embeddings," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2022, vol. 36, pp. 5521–5529.

[20] Shen Wang, Xiaokai Wei, Cicero Nogueira Nogueira dos Santos, Zhiguo Wang, Ramesh Nallapati, Andrew Arnold, Bing Xiang, Philip S Yu, and Isabel F Cruz, "Mixed-curvature multi-relational graph neural network for knowledge graph completion," in *Proceedings of the Web Conference 2021*, 2021, pp. 1761–1771.

[21] Ruobing Xie, Zhiyuan Liu, Huanbo Luan, and Maosong Sun, "Image-embodied knowledge representation learning," *arXiv preprint arXiv:1609.07028*, 2016.

[22] Zikang Wang, Linjing Li, Qiudan Li, and Daniel Zeng, "Multimodal data enhanced representation learning for knowledge graphs," in *2019 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2019, pp. 1–8.

[23] Meng Wang, Sen Wang, Han Yang, Zheng Zhang, Xi Chen, and Guilin Qi, "Is visual context really helpful for knowledge graph? a representation learning perspective," in *Proceedings of the 29th ACM International Conference on Multimedia*, 2021, pp. 2735–2743.

[24] Pouya Pezeshkpour, Liyan Chen, and Sameer Singh, "Embedding multimodal relational data for knowledge base completion," *arXiv preprint arXiv:1809.01341*, 2018.

[25] Zongsheng Cao, Qianqian Xu, Zhiyong Yang, Yuan He, Xiaochun Cao, and Qingming Huang, "Otkge: Multi-modal knowledge graph embeddings via optimal transport," *Advances in Neural Information Processing Systems*, vol. 35, pp. 39090–39102, 2022.

[26] Yu Zhao, Xiangrui Cai, Yike Wu, Haiwei Zhang, Ying Zhang, Guoqing Zhao, and Ning Jiang, "Mose: Modality split and ensemble for multimodal knowledge graph completion," *arXiv preprint arXiv:2210.08821*, 2022.

[27] Xinhang Li, Xiangyu Zhao, Jiaxing Xu, Yong Zhang, and Chunxiao Xing, "Imf: Interactive multimodal fusion model for link prediction," in *Proceedings of the ACM Web Conference 2023*, 2023, pp. 2572–2580.

[28] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[29] Xiang Chen, Ningyu Zhang, Lei Li, Shumin Deng, Chuanqi Tan, Changliang Xu, Fei Huang, Luo Si, and Huajun Chen, "Hybrid transformer with multi-level fusion for multimodal knowledge graph completion," in *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2022, pp. 904–915.

[30] Yinwei Wei, Xiang Wang, Liqiang Nie, Xiangnan He, Richang Hong, and Tat-Seng Chua, "Mmgcn: Multi-modal graph convolution network for personalized recommendation of micro-video," in *Proceedings of the 27th ACM international conference on multimedia*, 2019, pp. 1437–1445.

[31] Rui Sun, Xuezhi Cao, Yan Zhao, Junchen Wan, Kun Zhou, Fuzheng Zhang, Zhongyuan Wang, and Kai Zheng, "Multi-modal knowledge graphs for recommender systems," in *Proceedings of the 29th ACM international conference on information & knowledge management*, 2020, pp. 1405–1414.

[32] Valter Moretti, "The interplay of the polar decomposition theorem and the lorentz group," *arXiv preprint math-ph/0211047*, 2002.

[33] Alan Edelman, Tomás A Arias, and Steven T Smith, "The geometry of algorithms with orthogonality constraints," *SIAM journal on Matrix Analysis and Applications*, vol. 20, no. 2, pp. 303–353, 1998.