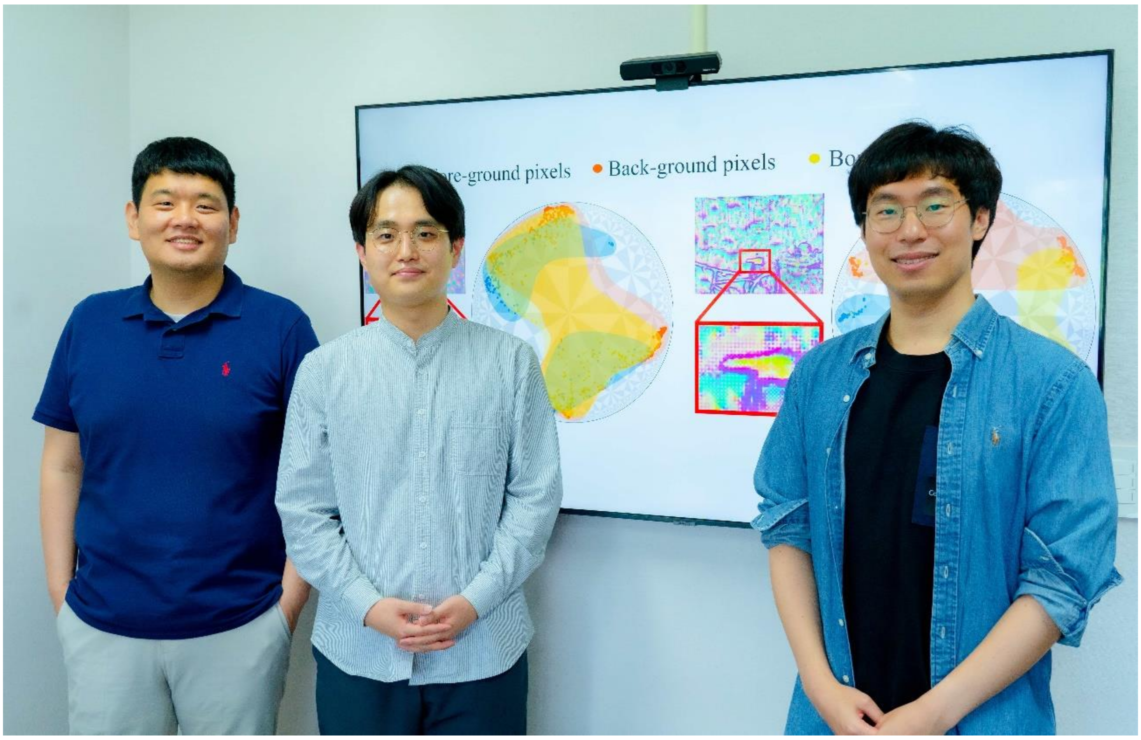
“Redefining Video Information in a Curved Space” GIST Develops AI Technology with Enhanced Image Processing Performance

- AI Graduate School Prof. Jeon Hae-gon's team presented on July 26th at the world’s top machine learning conference (ICML).
- Utilizing hyperbolic space... Learning affinity by efficiently redefining the relationship between pixels in an image.



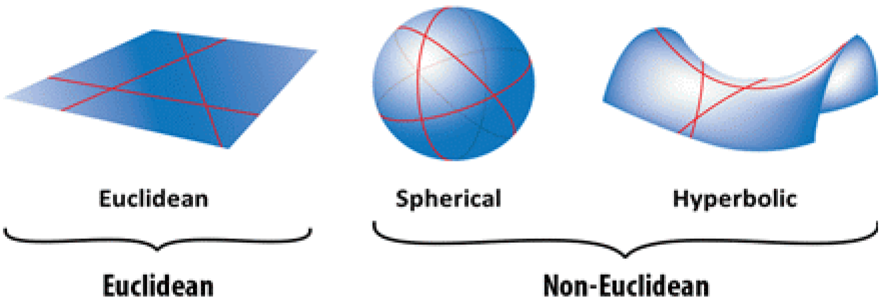
▲ (From left) Graduate School of AI Prof. Jeon Hae-gon, combined master's and doctoral student Park Jin-hui, and doctoral student Bae In-hwan

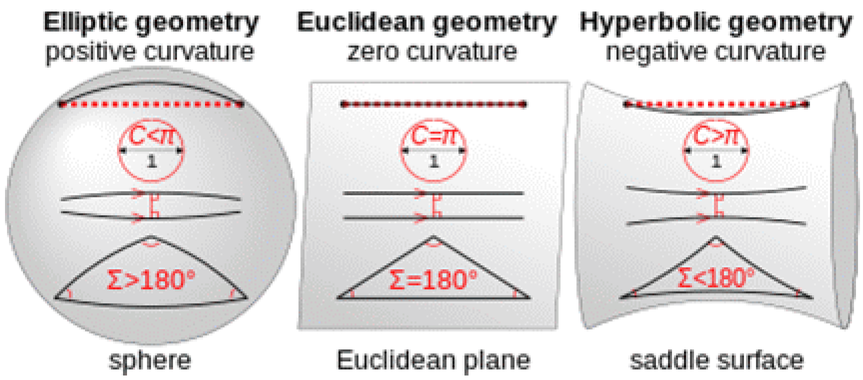
GIST (Gwangju Institute of Science and Technology, President Lim Kichul) has succeeded in developing artificial intelligence (AI) algorithm technology with greatly enhanced image processing performance by understanding the relationship between pixels in an image in a new way -- in a “hyperbolic space.”

Hyperbolic space is one of the non-Euclidean* spaces with a curved shape unlike the 3-dimensional space we experience on a daily basis. It is used to study various problems in mathematics and physics, and plays an important role in the theory of relativity, network structure research, and graph theory.

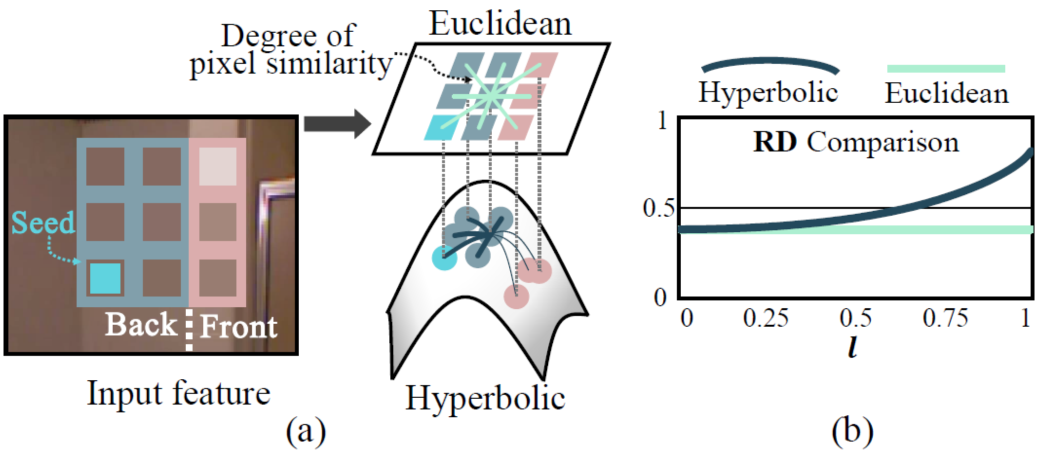
* Non-Euclidean Geometry: Geometry dealing with various spaces that modify or extend the curvature of space, the nature of parallel lines, and the definition of distance. Hyperbolic geometry is a representative example of non-Euclidean geometry and has negative curvature on a plane.

It has been proven that hierarchical relationships between data can be formed more efficiently in a hyperbolic space, and recently, it is more commonly applied in the AI fields in understanding the interrelationship of data.





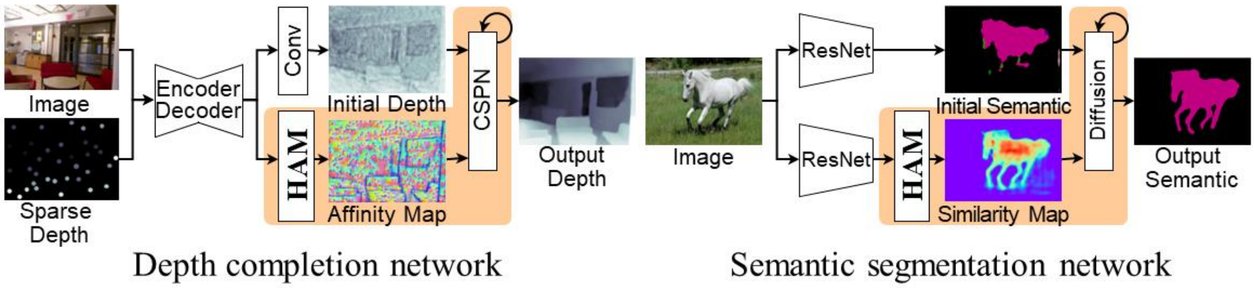
▲ Examples of Euclidean and non-Euclidean geometric spaces.



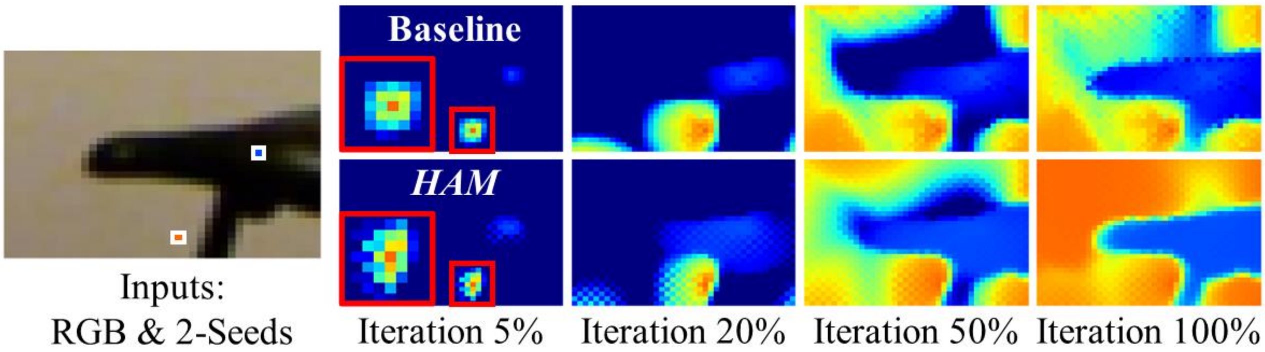
▲ (a) By redefining the pixel relationship in hyperbolic space, pixels that should be close can be placed closer and distant pixels farther. (b) Hyperbolic space is characterized by a larger distance relationship than Euclidean space.

GIST Graduate School of AI Prof. Jeon Hae-gon's research team developed a hyperbolic affinity learning module that understands affinity after layering the relationship between pixels in hyperbolic space, greatly improving image processing performance. It has been previously proven that layering the relationship between pixels in an image is efficient, but this is the first successful attmpt of this layering in hyperbolic space.

This module enhances the resolution of given information based on the information of neighboring pixels by using a spatial propagation algorithm that identifies the relationship between pixels and propagates the information. After inferring the affinity between pixels, sparse information present in low-resolution areas is diffused to the entire image.



▲ Hyperbolic spatial diffusion network structure proposed in the study. It spreads sparse information obtained from inexpensive sensors in image form to obtain precise results. The key feature of this study is that a more precise affinity map was obtained in the network by establishing the relationship between pixels in the hyperbolic space.



▲ An example of spatial diffusion results performed in this study. In the comparison of spatial diffusion results performed using images and two sparse information (leftmost), the method proposed by the researchers (HAM) shows more precise results than the method (Baseline) performed in the existing Euclidean space. In previous studies, a phenomenon in which information is leaked when spatial diffusion occurs due to the ambiguity of the boundary of an image object. The phenomenon was resolved using the hyperbolic affinity-learning module (HAM) proposed in this study, and more precise results were obtained.

In the curved hyperbolic space, close objects are expressed as being closer and distant objects as being farther away, sallowing efficient extraction of image features. In this manner, the ability to accurately recognize the boundaries between objects was increased by up to 14%.

The research team also proposed a new method of arranging learning features based on geodesic* distance, which is the shortest distance in hyperbolic space.

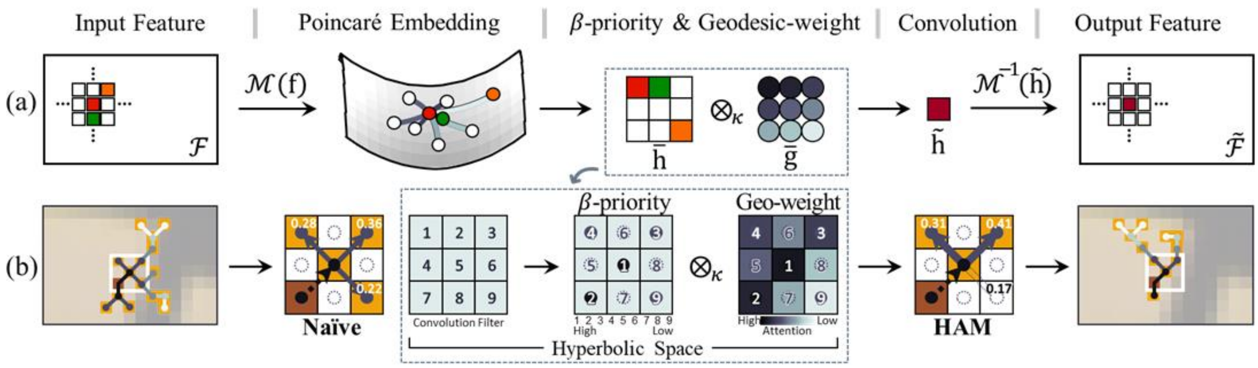
Because locational information may be lost in the process of converting to hyperbolic space, newly expressed features were arranged in the order of geodesic distance and given locational information again. The closer the distance, the higher the affinity, suggesting that the objects are identical.

* Geodesic: A curve or space representing the shortest path between two points, in a space with curvature, it appears bent or twisted.

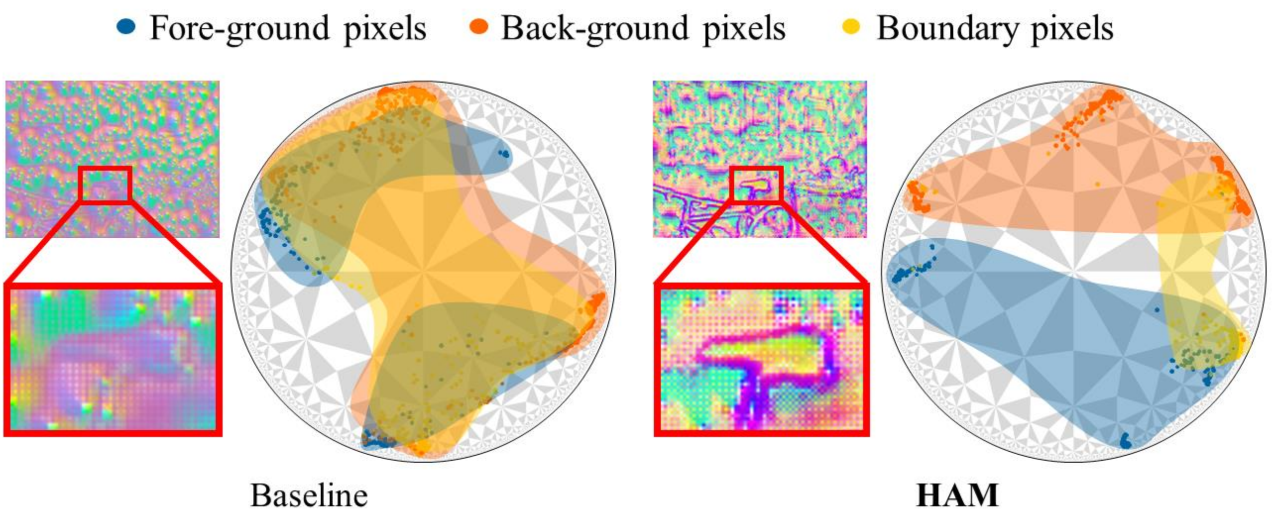
In particular, the methodology proposed by the research team was designed as a plug-and-play* module to allow easy application to existing algorithms that infer or semantically segment 3D depth information of images.

* Semantic Segmentation: This refers to dividing images into pixels, analyzing which object each pixel belongs to, and classifying them as roads, cars, pedestrians, trees, etc.

* Plug-and-Play: A function that allows one to use any device without any additional physical setting other than installation.



▲ The hyperbolic affinity learning module (HAM) proposed in this study. In this study, we propose a module that can learn and infer the affinity between pixels in hyperbolic space. It has been demonstrated in several computer vision fields that layering the relationship between pixels is efficient. This study allows this pixel layering to be performed in hyperbolic space. For effective learning, priorities are assigned based on hyperbolic geodesic distances, and weights are assigned.



▲ Comparison of affinity map visualization between this study and previous studies. On the left (baseline) is the result of inferring the pixel affinity map in Euclidean space performed in previous studies, and on the right (HAM) is the hyperbolic affinity result inferring with the methodology proposed by the researchers. A clearer distinction between objects can be observed.

Prof. Jeon Hae-gon stated, "Through this research, it is now possible to improve applied algorithms for images and 3D space in the field of computer vision and robotics without additional learning parameters." He went on to say, "We expect to be able to expand on this in a wider field of research, such as the relationship between objects in images or 3D space, going beyond inter-pixel relationships."

This research, led by Prof. Jeon Hae-gon, conducted by combined master's and doctoral student Park Jin-hwi, and participated in by doctoral student Bae In-hwan and others, was supported by the GIST-MIT collaborative research project and the AI Innovation Hub project of the Ministry of Science and ICT. This research is scheduled to be presented on July 26 at the International Conference on Machine Learning (ICML), the world's top conference in artificial intelligence and machine learning.

GIST
Since 1993