

测地空间中泊松碟采样的地形模型约简方法

侯文广, 吴梓翠, 丁明跃

(华中科技大学生命科学与技术学院图像信息处理与智能控制教育部重点实验室, 湖北武汉 430074)

摘要: 针对地形模型的数据冗余问题, 提出在测地空间中进行均匀采样的策略, 实现数据精简的目的. 该方法随机地从原始地面模型中提取点, 根据当前点与已有采样点的测地距离确定是否将其加入到采样点集, 重复该过程直至无新的采样点加入. 通过定义加权测地距离, 该方法能有效调整不同区域的采样概率, 从而使精简后的模型能有效保持原始数据的特征信息. 该算法原理简单, 实现容易, 约简后数据的分布具有良好的可视化效果.

关键词: 泊松碟; 地形模型约简; 测地距离; 自适应采样

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112 (2012) 06-1274-05

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2012.06.035

Poisson Disk Sampling in Geodesic Metric for DEM Simplification

HOU Wen-guang, WU Zi-cui, DING Ming-yue

(College of Life Science and Technology, Huazhong University of Science and Technology, Image Processing and Intelligent Control Key Laboratory of Education Ministry of China, Wuhan, Hubei 430074, China)

Abstract: A geometric approach for DEMs simplification is proposed. The main idea is to uniformly pick points from DEM in the sense of geodesic metric, resulting in terrain-adaptive samples in Euclidean metric. This method randomly sample point from mesh nodes, then judge if the point can be accepted or not according to its geodesic distance from sampled point set, and repeat the whole process until no points can be added. By defining weighted geodesic distance positively related to terrain variation, results with more points in rugged terrain areas and sparse points in flat areas can be obtained. Moreover, the distribution of samples is very fit for high performance terrain visualization. This method is really simple and can retain the topographical details more effectively.

Key words: Poisson disk; DEM simplification; geodesic distance; adaptive sampling

1 引言

数字地形模型(Digital Elevation Model, DEM)的精简广泛应用于大数据量的地形场景绘制、地形模型 LOD 生成和地图综合等应用中, 目的是在尽可能保持地形特征信息的基础上减少数据冗余, 期望的结果是在地形复杂区域采样比较密集, 反之亦然. 地形模型的精简是测绘领域的经典问题, 许多学者针对该问题展开了大量深入的研究^[1-3], 本文将主要讨论 DEM 的几何压缩方法, 且仅考虑由格网或不规则三角网表示的情况.

建立金字塔模型是一种相对简单的数据约简方案, 缺点是不能有效保持地形的特征信息, 而且采样点集的规则性会产生混沌现象^[1]. 与格网相比较, 不规则三角网能有效保持地形的特征信息并显著压缩数据量, 故利用顶点、鞍部点和拐点等特征点建立不规则三角网模型成为了研究的重点, 文献[4, 5]提出在不同尺度空间中利用聚类算法提取特征点; 文献[6]在尺度空间理论和 Voronoi 图的基础上进行特征提取; 启发式算法通过不

断迭代寻找最优的采样点集. 文献[7]提出一种格网和不规则三角网联合表示 DEM 的方法, 有效地融合二者的优势. 一些学者采用小波、分形等方法进行 DEM 的精简, 取得了一定的进展^[8]. 考虑到初始生成的不规则三角网还可能存在一定的数据冗余, Hoppe 较早地提出渐进的几何约简方法, 提供了不规则三角网几何模型约简的通用算法^[9]. 文献[10]依据相邻三角形法向量的夹角, 通过合并相邻的三角形实现压缩的目的. 若将 DEM 的高程视作是图像的灰度, DEM 的平面格网作为图像的像素坐标, 则对 DEM 的处理可以转换为图像处理有关问题. 文献[11]在融合了图像几何表示和无损压缩方法的基础上, 提出了一种精简方法; 文献[12]比较全面总结和分析了图像的几何表示方法在 DEM 数据处理中的应用. 这些方法的缺点是算法相对复杂, 结果的自适应性不够.

在图像处理领域, 广泛采用的均匀采样方法是泊松碟(Poisson disk)和最远点策略(Farthest Point Strategy, FPS), 在实际应用中, 它们不足之处是采样过程没有考

虑原始图像的信息,无法直接得到自适应分布的结果.文献[13]提出一种利用初始均匀采样点获得自适应结果的方法,但算法的灵活性不够.为了实现自适应采样的目的,本文认为应将数据放置于一个非欧氏的空间,该空间能够反映局部图像区域的变化.然后,在非欧氏空间中进行均匀采样后,能得到在欧氏空间中自适应分布的结果.基于该思想,并受计算机图形学中测地空间的启发,本文提出了在测地空间中进行均匀采样的方法,将其应用于 DEM 的压缩中.测地空间的基本特性是:如果具有一定欧氏距离的两个点位于地形平坦的区域,则其测地距离接近于欧氏距离;如果他们位于地形变化大的区域,测地距离应该显著大于其欧氏距离.因此,在测地空间中均匀采样,地形复杂区域的点被采样的概率增加,而平坦区域采样机会相对较少,从而实现自适应采样的目的.原理如图 1 所示, A 、 B 和 C 均匀分布于二维的曲线上,但 AB 之间的欧氏距离大于 BC ,因为 AB 间的形状变化相对较小.从理论和实际上,FPS 和泊松碟采样将得到同等质量的采样点,但是前者的计算量较大,故本文仅采用泊松碟采样.

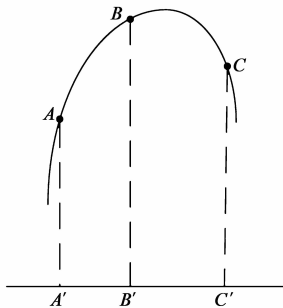


图1 测地距离

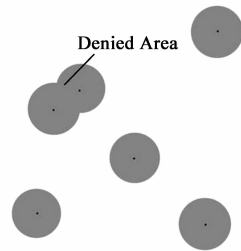


图2 泊松碟采样原理

2 泊松碟采样

在一定的空间中,泊松碟采样是要得到这样一群点集:每个采样点是随机生成的;任意两个采样点之间的间隔要大于给定的阈值;在该空间内无新的点可插入^[14,15].在平面中,采样点的分布如图 2 所示,新的采样点只能随机地出现在图中的白色区域,得到新采样点后,以它为中心的一定区域将成为后续采样的禁区,按此准则采样直到整个平面上没有白色区域后,得到的采样点就满足泊松碟条件.显然,在各向同性的欧氏空间中,该方法将得到分布均匀但不规则的采样点,采样点的这种分布能有效避免图像重建所出现的混沌现象,正是该特征使得其在图像采样中具有重要的价值^[13].

Cook 首先提出了实现泊松碟采样的经典方法投掷算法^[14],其原理是:在给定空间中随机投掷(生成)一个采样点,计算其与已有采样点集合中所有点的距离,

若不存在小于阈值的情况,则将新的点加入到采样点集中,否则,重复以上的过程直至无新的采样点加入为止.投掷算法是泊松碟采样的严格方法,缺点是在需求大量采样点时,计算相当耗时.但在本文中,泊松碟采样将用于从格网模型中随机采集点,与常规投掷法中生成采样点的过程略有不同;另一不同就是该采样是在 DEM 的测地空间中,该空间不是各向同性的,很多快速算法无法直接应用,故本文仅考虑经典的投掷算法.

3 测地距离

如图 1 所示,平面曲线上有三个点 A 、 B 和 C ,沿着曲线从 A 到 B 或 C 的最短弧长为对应的测地距离.于是,测地距离与欧式距离的差反映了局部形状的变化量,故在测地空间中均匀采样后,形状变化大的区域采样密集,而平坦区域采样相对稀疏.相应地,对于二维流形曲面,任意两点的测地距离为沿曲面表面的最短弧长.从理论上而言,测地距离的计算只有在参数曲面模型条件下才能进行;对于由网格所表示的曲面而言,测地距离通常由以下算法近似估计:在欧氏距离下一定邻域内的点,测地距离等于欧氏距离;对于不在邻域内的点,其测地距离由连接邻域内点的最短路径逼近.在本文算法实现中,Fast Marching 用于测地距离的计算^[16].

在某些应用中,通常测地距离定义使得采样点集的自适应程度不够.为了能够灵活有效调整采样点的分布,本文定义了加权的测地距离,公式(1)就是一个简单的加权测地距离的计算方法.

$$d_s = |X_p - X_{p_0}| + |Y_p - Y_{p_0}| + k \times |H_p - H_{p_0}| \quad (1)$$

d_s 表示点 P 与其邻域点 P_0 的测地距离, (X_p, Y_p, H_p) , $(X_{p_0}, Y_{p_0}, H_{p_0})$ 分别表示 P 和 P_0 的空间坐标, $k > 1$ 将使地形复杂区域点的加权测地距离显著增加,从而提高地形复杂区域的采样概率.当然,这仅仅是一个简单的测地距离估算的例子,更加复杂的方法不在此讨论.与该加权测地距离的定义相对应,非邻域内任意两点的加权测地距离,由连接邻域内点的最短路径的加权测地距离之和逼近.

4 算法描述

设 $Disk$ 为给定的测地距离的阈值,采样点的集合为 S ,基于泊松碟原理,对 DEM 的约简过程为:

步骤 1 $S = \emptyset$,待采样 DEM 的三维点的 $Sampled$ 和 $Visited$ 属性为 $false$;

步骤 2 随机地从三维点集中选取一个点,设为 $P(x, y, z)$;

步骤 3 若该采样点的 $Sampled$ 或 $Visited$ 的属性为

true,增加无效投掷次数,若无效投掷次数大于给定的阈值,终止;否则,返回步骤 2;

步骤 4 令无效投掷次数为零;

步骤 5 将 P 从加入到 S ,从点 P 开始,在 DEM 中计算到点 P 测地距离小于 Disk 半径的点集,设为 V (不包含点 P 本身);

步骤 6 标记 P 的 Sampled 属性为 true,标记 V 中点的 Visited 属性为 true;

步骤 7 返回步骤 2.

步骤 5 将采用 Fast Marching 方法实现.在网格上行进时,若沿任意边的行进速度相同,则计算距离某个点一定测地距离内的点等同于搜索一定时间内所能达到的全部点.以格网为例,如图 3 所示,设从网格中的 A^0 开始,显然在相同时间内可以同时达到 A^1 ,标记 A^0 为 Reached, A^1 为 Alive,随后的计算过程如下:

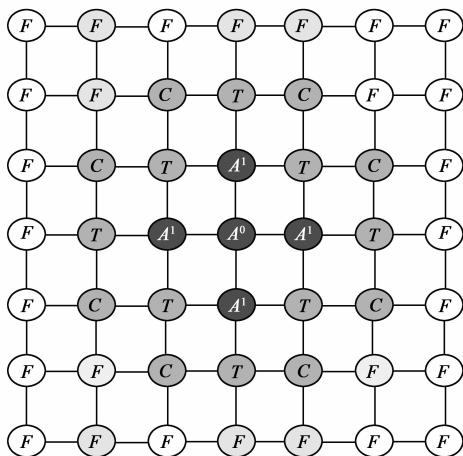


图3 Fast Marching

步骤 1 标记 Alive 近邻域内非 Reached 的点为 Close,其它点为 Far(在 8 近邻的情况下,图 3 中 C 和 T 节点都属于 Close; F 节点为 Far.在三角网格中,Alive 近邻域内的点是指与其直接相连的点);

步骤 2 计算 Close 集合中点的达到时间,若某 Close 点可从不同线路到达,取最小达到时间;

步骤 3 具有最小到达时间或到达时间差小于一定范围的点标记为 Trial,对应于图 3 中的 T ,将这部分 Trial 插入到 Alive 集合中;

步骤 4 若某点邻域内点均为 Alive,标记该点为 Reached;

步骤 5 若 Alive 点的到达时间大于等于给定的阈值,输出 Alive 和 Reached 中的点,终止循环;

步骤 6 返回步骤 1.

从以上算法可知,若不考虑 DEM 的高度变化,同样的测地距离在不同区域都将覆盖相同数量的点(边缘除外),故采样结果将是均匀的.若考虑 DEM 的高度变

化,与地形平坦区域相比较,在地形复杂区域,同样的测地距离将覆盖较少的 DEM 点.由于投掷过程是随机且均匀的,故在地形复杂区域将会有更多的机会得到采样点,从而使采样结果能保持原始地形的特征信息.

5 参数设置

在以公式(1)作为计算邻域点加权测地距离的情况下,为了能够在地形复杂区域得到更多的采样点, k 显然应大于 1,但是若 k 过大,将导致过多的点集中于地形变化区域,同时也使平坦区域的采样点过于稀疏.在通常的应用中, $k \in [2,4]$ 是一个较为合适的范围.

另一个关键的参数就是泊松碟的采样半径,在测地距离计算方法确定的条件下,该半径值将基本决定采样点的数量(从泊松碟原理可知,不能在采样前准确确定采样点的数量).假如原始 DEM 点数为 N ,期望采样点的数量 N_s ,计算半径值的方法如下:任意从原始数据中选择一个点,从该点出发,采用 Fast Marching 方法从该点向外扩张计算到其它点的测地距离,当访问点的数量为 N/N_s 时,获得已访问点中的最大测地距离,然后再随机从原始数据中选择点,重复上述的过程若干次(20 次左右),将最大测地距离的平均值作为泊松碟的半径.

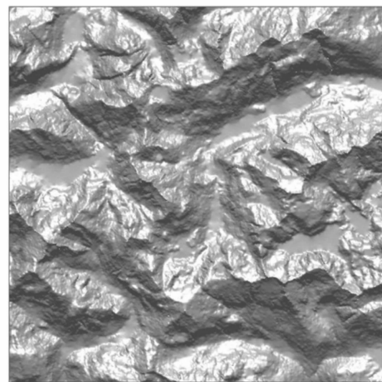


图4 原始DEM模型

6 实验与分析

本实验数据为某地区的数字地面模型,分辨率为 20m,原始数据由规则格网表示,格网大小为 512×512 ,如图 4 所示,地形变化复杂.首先,按照规则格网对 DEM 进行采样,结果如图 5 所示,出现了明显的混沌现象,可视化效果比较差.按照泊松碟采样法在 DEM 的平面网格中采样,如图 6 所示,该方式没有考虑到地形的变化信息,但采样点的分布是非规则的,但其可视化的效果明显优于规则采样的结果.采用本文提出的自适应采样方法,按照公式(1)计算格网点 8 邻域点的测地距离,令 $k = 4$,采用 Fast Marching 搜索测地距离为碟半径范围内的点,采样结果如图 7 所示,生成图 7(a)和 7(b)时,碟半径分别为 420m 和 90m.显然,点云分布与地形的变化相互适应,并且点云没有过渡集中在地形变化的区域.从中可知,本文算法实际是在寻找自适应采样和均匀采样的平衡点,采样点集不但考虑到地形的

变化,而且其分布适合高质量的三维显示.对于该地形数据,我们在其不规则三角网模型中进行了自适应采样,在相同的条件下,能够得到类似的结果,本文不再列举.图 8 是利用其它三维处理软件采样的结果,8(a)是利用 Geomagic 中的“曲率采样”的结果;而 8(b)是利用 ArcGIS 中“Subset”计算得到.比较图 7 和图 8,本文自适应采样算法表现出更好的能力.本实验中,所有三角网格均利用三维点的平面坐标,按照二维 Delaunay 准则构建得到.

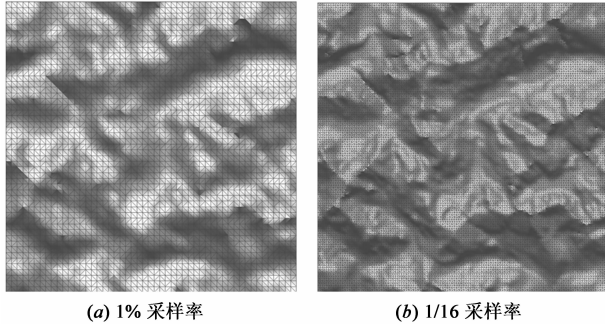


图5 规则采样结果

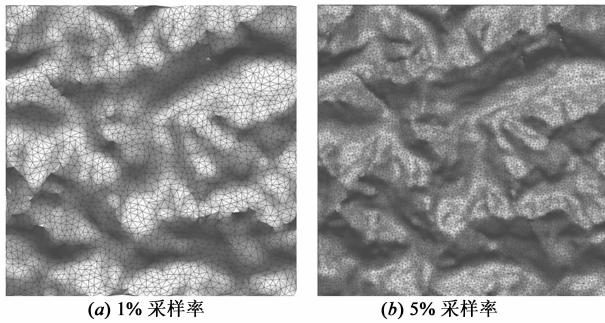


图6 均匀非规则采样结果

图 9 是采用 4 近邻内插方法对图 6(a)和图 7(a)重建后的 DEM,计算公式如(2)所示.

$$H(p) = \sum_{i=1}^4 \frac{H(s_i)}{d(p, s_i)} / \sum_{i=1}^4 \frac{1}{d(p, s_i)}, \quad (2)$$

其中, $\{s_i\}_{i=1}^4$ 是点 p 的四个近邻点, $H(p)$, $H(s_i)$ 分别为点 p , s_i 的高程, $d(p, s_i)$ 为 p , s_i 之间的距离.从视觉效果看,图 9(b)的效果要优于图 9(a).为了对约简后的模型进行精度评价,采取如下的策略:首先构建约简后点集的三角网格模型;根据原始 DEM 点的平面坐标判断该点位于约简后网格模型的那个三角形中;利用三角形的顶点坐标内插得到其对应的高度值;计算内插的高度和原始高度的差.该方法相对简单,但能比较客观反映约简后模型的整体精度.在本实验中,按照上述方法计算的中误差如图 10 所示.显然,随着采样率的提高,误差明显减小;在相同采样率的情况下,自适应采样的结果要优于均匀采样的结果;当采样率达到 50% 时,约简后模型的误差约为 0.3m,精度损失有限.

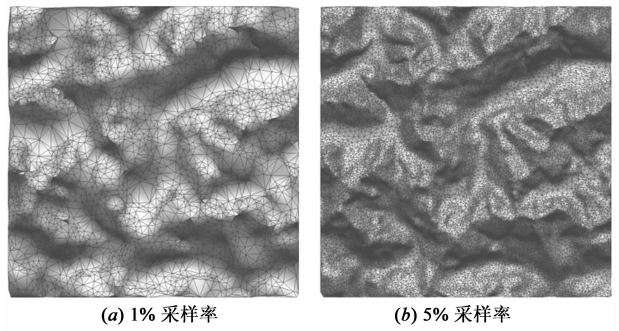


图7 自适应采样结果

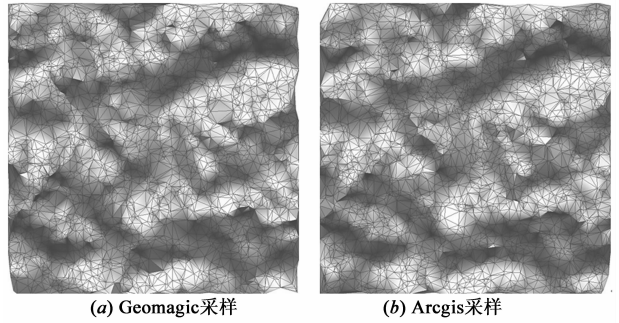


图8 Geomagic, Arcgis 的采样结果,采样率为1%

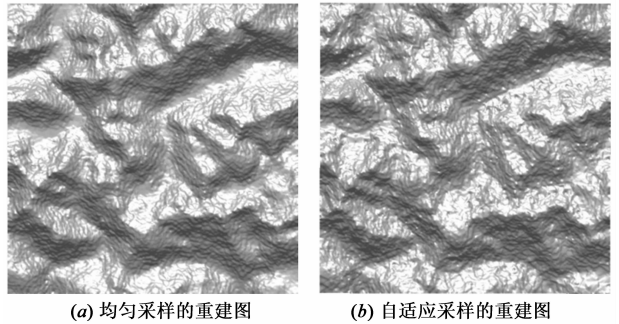


图9 图6(a)和图7(a)重建的地面模型

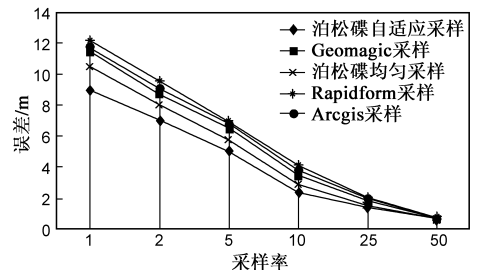


图10 重建模型与原始DEM的误差

7 总结

基于图像处理中泊松碟采样的基本原理和计算机图形学中的测地空间的理论,本文提出一种在格网或不规则三角网模型中自适应采样的方法,并将其应用于 DEM 的几何压缩.该方法具有如下优点:采样点集的分布具有自适应性,约简后模型能较好地保持原始数据的特征信息;采样点的分布兼顾了地形变化和分布的均匀性,使得采样结果特别适合三维可视化.进一步

的工作将研究测地距离的估算方法,以及如何自适应地确定采样半径等.

参考文献

- [1] Brassel K, Weibel R. A review and conceptual framework of automated map generalization[J]. *International Journal of Geographic Information System*, 1988, 2(3): 229 – 244.
- [2] López A, Lumbreras F, Serrat J, et al. Evaluation of methods for ridge and valley detection[J]. *IEEE Trans PAMI*, 1999, 21(4): 327 – 335.
- [3] Ai T, Li J. A DEM generalization by minor valley branch detection and grid filling[J]. *ISPRS Journal of Photogrammetry & Remote Sensing*, 2010, 65(2): 198 – 207.
- [4] Song H, Feng H. A global clustering approach to point cloud simplification with a specified data reduction ratio[J]. *Computer-aided Design*, 2008, 40(3): 281 – 292.
- [5] 郭庆胜, 郑春燕, 胡华科. 基于邻近图的点群层次聚类方法的研究[J]. *测绘学报*. 2008, 37(2): 255 – 261.
Guo Q, Zheng C, Hu H. Hierarchical clustering method of group of points based on the neighborhood graph[J]. *Acta Geodaetica ET Cartographica Sinica*, 2008, 37(2): 255 – 261. (in Chinese)
- [6] Ali T. Building of robust multi-scale representations of lidar-based digital terrain model based on scale-space theory[J]. *Optics and Lasers in Engineering*, 2010, 48(3): 316 – 319.
- [7] Yang B, Shi W, Li Q. An integrated TIN and grid method for constructing multi-resolution digital terrain models[J]. *International Journal of Geographical Information Science*, 2005, 19(10): 1019 – 1038.
- [8] Wu F. Scale-dependent representations of relief based on wavelet analysis[J]. *Geo-Spatial Information Science*, 2003, 6(1): 66 – 69.
- [9] Popovic J, Hoppe H. Progressive simplicial complexes[A]. *ACM SIGGRAPH 1997 Conference Proceedings [C]*. New York: ACM Press, 1997. 217 – 224.
- [10] 刘春, 吴杭彬. 基于平面不规则三角网的 DEM 数据压缩与质量分析[J]. *中国图象图形学报*, 2007, 12(5): 836 – 840.

Liu C, Wu K. Data compress of DEM based on plane triangulation irregular network[J]. *Journal of Image and Graphics*, 2007, 12(5): 836 – 840. (in Chinese)

- [11] Sole A, Caselles V, Sapiro G, et al. Morse description and geometric encoding of digital elevation maps[J]. *IEEE Trans on Image Processing*, 2004, 13(9): 1245 – 1262.
- [12] Caselles V, Monasse P. Geometric Description of Topographic Maps and Applications to Image Processing[M]. Heidelberg, London: Springer, 2008.
- [13] Eldar Y, Lindenbaum M, Porat M. The farthest point strategy for progressive image sampling[J]. *IEEE Trans on Image Processing*, 1997, 6(3): 1305 – 1315.
- [14] Cook R. Stochastic sampling in computer graphics[J]. *ACM Trans Graph*, 1986, 5(1): 137 – 145.
- [15] Moenning C, Dodgson N. Fast Marching Farthest Point Sampling[R]. Cambridge: University of Cambridge, UK, 2003.
- [16] 陈松, 王珊, 周明天. 基于实时分析的网络测量抽样统计模型[J]. *电子学报*, 2010, 38(5): 1177 – 1180.
Chen S, Wang S, Zhou M. Network data measurement and statistics model based on real time analysis[J]. *Acta Electronica Sinica*, 2010, 38(5): 1177 – 1180. (in Chinese)

作者简介



侯文广 男, 1974 年生, 华中科技大学生命学院教师, 主要研究方向为图像处理和计算机视觉.

E-mail: houwenguang99@163.com

吴梓翠 女, 1984 年生, 华中科技大学生命学院硕士生, 主要研究方向为图像处理.

丁明跃 男, 1961 年 5 月生于湖北麻城, 现为华中科技大学特聘教授, 博士生导师, 长期从事图象处理、模式识别和计算机视觉等领域的科研工作.