

DOI: 10.3724/SP.J.1187.2012.00138

三维人脸深度图的流形学习-LOGMAP 识别方法*

詹 曙¹ 张芝华¹ 叶长明¹ 蒋建国^{1,2} S.Ando³

(1. 合肥工业大学计算机与信息学院, 合肥 230009; 2. 安全关键工业测控技术教育部工程研究中心, 合肥 230009; 3. 日本东京大学信息科学学院, 东京 113-8656)

摘 要: 人脸识别是生物特征识别技术最友好的身份识别方式, 而三维人脸识别由于可解决二维人脸识别中存在的光照、姿态等局限, 成为人脸识别的研究热点, 但其特征维数过高是该领域的瓶颈, 而维数约减是解决这一问题的关键。流形学习是一类非线性维数约减算法, LOGMAP 是一种基于黎曼法坐标的流形学习算法, 该算法可以在保持度量信息不变的情况下, 把高维空间的数据映射到低维空间。构建了基于流形学习的三维人脸识别框架, 并结合 LOGMAP 进行三维人脸深度图的识别。实验结果表明, 该方法在三维深度图上可以得到良好的识别效果。

关键词: 三维深度图; 人脸识别框架; 流形学习; 黎曼法坐标; LOGMAP

中图分类号: TP391

文献标识码: A

国家标准学科分类代码: 510.4050

3D facial depth map recognition based on manifold learning-LOGMAP algorithm

Zhan Shu¹ Zhang Zhihua¹ Ye Changming¹ Jiang Jianguo^{1,2} S. Ando³

(1. School of Computer & Information, Hefei University of Technology, Hefei 230009, China; 2. Engineering Research Center of Safety Critical Industrial Measurement and Control Technology, Ministry of Education, Hefei 230009, China; 3. Department of Information Physics and Computing, the University of Tokyo, Tokyo 113-8656, Japan)

Abstract: Face recognition is the no-touch authentication technology in biometrics. 3D face recognition becomes more popular because it can solve the problems of the effect from different illumination and pose in 2D face recognition. However, how to decrease the number of feature dimension is a key technology in 3D facial recognition. In this paper, LOGMAP manifold learning algorithm based on normal coordinates recently is proposed to establish the mapping relationship between the observed and the corresponding low-dimensional data. This paper built a 3D face recognition framework based on manifold learning, and applied LOGMAP theory into 3D facial depth image recognition. The experimental results demonstrate that this method can get good effect for recognition.

Keywords: 3D facial depth map; face recognition framework; manifold learning; Riemannian normal coordinates; LOGMAP

1 引 言

随着计算机技术和生物医学工程技术突飞猛进的发展, 利用人体生物特征来鉴别个体身份的生物学识别技术已成为安全验证的首选方式。人脸识别技术以其所具有的非接触、非侵犯和无排斥等优势, 成为最具发展潜力的身份验证手段之一, 引起研究者的广泛关注。

目前基于二维图像的人脸识别技术已日渐成熟, 但二维人脸识别技术在光照、姿态和有遮挡等非理

想采样下得不到理想的识别效果, 鲁棒性较差。三维人脸数据以其丰富的信息量、光照不变和姿态不变等优点解决了二维人脸数据存在的光照、姿态和表情这三大瓶颈问题^[1], 使得越来越多的学者转向三维人脸识别^[2]的研究。但三维人脸数据特征维数过高又阻碍了其发展, 维数约减是解决这一问题的关键。

在模式识别领域, 维数约减已成为一个重要研究方向, 其经典算法有主成分分析 PCA^[3]和多尺度化 MDS^[4], 但这两种算法仅适于处理欧式数据, 对

本文于 2011 年 10 月收到。

*基金项目: 国家自然科学基金(61174170); 教育部博士点基金(20101111110005)资助项目。

于非欧式数据的维数约减则不适用。为了适于非欧式数据的处理,近年来提出了大量的流形学习算法,包括等距特征映射 ISOMAP^[5-6]、局部线性嵌入 LLE^[7-8]、拉普拉斯特征映射 LE^[9]、局部切线空间排列 LTSA^[10]。这些维数约减算法可分为全局算法和局部算法两大类,如图1所示。

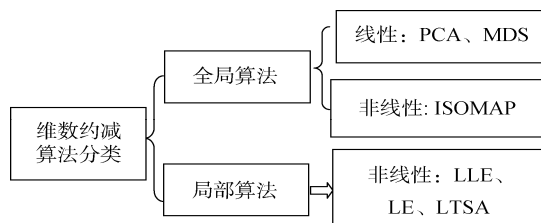


图1 维数约减算法分类

Fig. 1 Classification of dimension reduction algorithm

LOGMAP^[11]是 A.Brun 等提出的一种基于黎曼法坐标的非线性维数约减方法,该方法在保持数据本质特征的情况下把高维空间的数据映射到低维空间。本文构建了基于流形学习的三维人脸识别框架,并结合 LOGMAP 对三维深度图识别^[12]。实验结果表明,该方法可以区别不同人脸之间的差别,识别效果良好。

2 数学基础

2.1 黎曼流形定义^[13]

流形: 设 M 是一个 Hausdorff(豪斯道夫)空间。若对空间任意一点 $x \in M$, 都有 x 在 M 中的一个邻域 U 同胚与 m 维欧式空间 R^m 的一个开集, 则称 M 是一个 m 维流形(或拓扑流形)。

黎曼流形: 若在 m 维光滑流形 M 上给定一个光滑的处处非退化的对称二阶协变量场 G , 则称 M 是广义黎曼流形。当 G 是正定时, 称 M 是黎曼流形。

2.2 流形学习定义^[13]

假设数据 $X(x_1, x_2, \dots, x_n)$, $x_i \in R^D$ 是均匀采样与一个高维空间 M 的样本值, 流形学习就是从该高维数据探测出低维流形 $Y(y_1, y_2, \dots, y_n)$, $y_i \in R^d$, 其中 $d \ll D$, 即找到高维空间中的内蕴流形, 并求出相应的嵌入映射。

2.3 黎曼法坐标^[11,13]

流形学习从高维到低维映射的关键是要尽可能多的保留数据的度量信息, 而黎曼法坐标的优点是:

在不可能保持所有度量的情况下, 可以在某种程度上保持测地线度量, 其几何解释如图2所示^[9], 其中 p 是一特定点(基点), a, b, c 是球面上的任意三点, 根据 p 到这三点的距离和方向信息, 把这三点用对数映射(LOGMAP)映射到二维平面上, 得到点 a', b', c' , 这三点就是 a, b, c 在低维空间的黎曼法坐标。

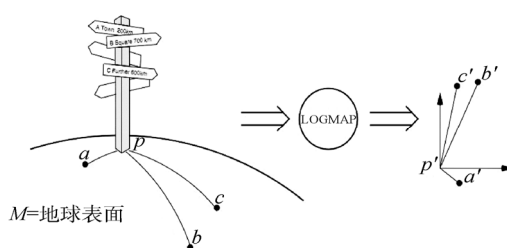


图2 黎曼法坐标

Fig. 2 Riemannian normal coordinates

3 LOGMAP 在三维人脸识别中的应用

3.1 三维人脸识别框架

三维人脸特征维数太高, 以至于无法进行有效的分析和识别, 所以在识别前我们对原有的高维数据进行维数约减。

人脸有基本的面部特征(如眼睛、鼻子、嘴巴)和相似的面部动作(如眼睛的睁闭、嘴巴的开合), 即不同人脸数据都有相似的特征, 这样不同的人脸图像可认为有相同的感知概念^[14]。因此我们可以把不同人脸数据映射到同一个子流形^[15], 进行维数约减。然而人脸除了公共特征外还有个性特征, 比如眼睛的大小、眉毛的粗细和鼻子的高低等, 我们可以利用这些个性特征在低维空间进行识别、区分不同的人脸。其系统识别框架如图3所示。

3.2 典型维数约减算法在三维深度图应用

3.2.1 主成分分析 PCA^[3]

PCA 是模式识别中最常用的一种线性维数约减算法, 其基本思想是: 在保持数据绝大部分信息的情况下, 把高维空间的数据映射到低维空间, 达到维数约减的目的。但该算法是线性降维算法, 而三维深度图是典型的非线性流形, 所以 PCA 不适于三维人脸的识别。 $d=3$, PCA 对三维人脸深度图降维如图4所示。

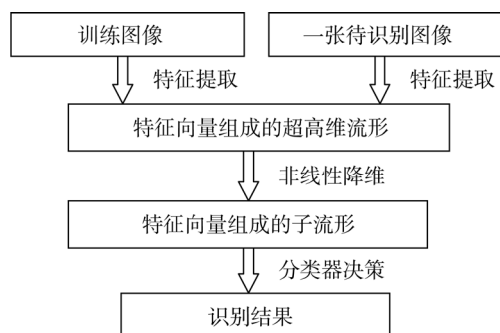


图 3 三维人脸识别框架

Fig. 3 3D face recognition framework

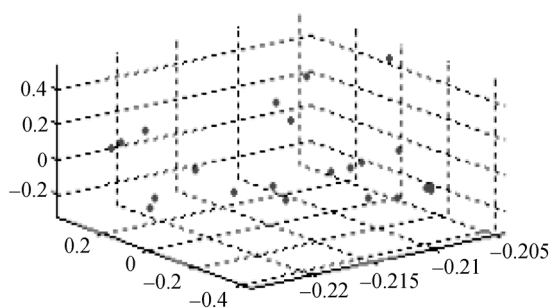


图 4 基于 PCA 的三维深度图维数约减

Fig. 4 Dimension reduction of 3D depth image using PCA

3.2.2 等距特征映射 ISOMAP^[5-6]

ISOMAP 是一种典型的全局非线性优化算法，建立在 MDS 基础之上。其主要思想是：在高维空间与低维空间之间建立等距特征映射，保持测地线距离不变，即保留了数据固有的非线性几何结构。但 ISOMAP 仅适于零高斯曲率的流形，不适于处理有较大内在曲率的流形，且计算量比较大，所以 ISOMAP 不适于三维人脸识别。 $d=3$ 、 $k=14$ ，ISOMAP 对三维深度图维数约减如图 5 所示。

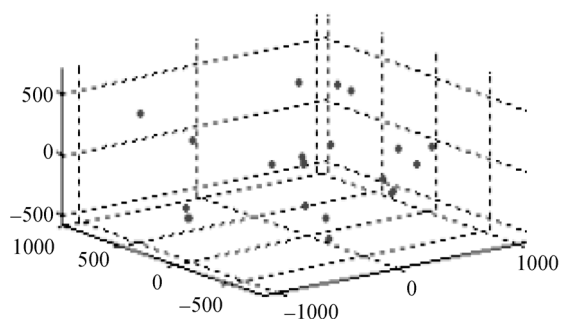


图 5 基于 ISOMAP 的三维深度图维数约减

Fig. 5 Dimension reduction of 3D depth image using ISOMAP

3.2.3 局部线性嵌入 LLE^[7-8]

LLE 是一种典型的局部非线性流形学习算法。其基本思想是：流形是局部线性的，那么在局部邻域，一个样本点就可以由它的邻域点线性重构而成，而且重构权值使得样本点与邻域点的线性重构误差达到最小。对每个样本点都重复以上最优重构权值的求解，这样可以保持嵌入空间与本质空间中数据的局部关系不变，得到高维空间在低维空间的真实嵌入。 $d=3$ 、 $k=8$ ，LLE 对三维深度图维数约减如图 6 所示。

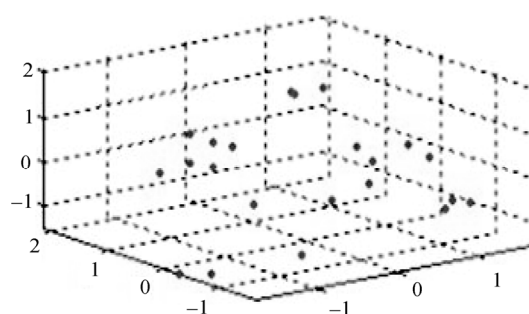


图 6 基于 LLE 的三维深度图维数约减

Fig. 6 Dimension reduction of 3D depth image using LLE

3.2.4 LOGMAP^[11-12]

LOGMAP 是基于黎曼法坐标提出的非线性流形学习算法，其基本思想是：用低维空间的黎曼法坐标来表示高维数据集上的点，达到维数约减的目的。其做法是在数据集 X 上选择一个基点 p ，计算其 k 最近邻表示为 $B(p)$ ，在 $B(p)$ 的外空间执行 PCA，得 p 点的切空间 T_pM ，在 T_pM 上取一组标准正交基 $(p; e_1, e_2, \dots, e_d)$ ，以这组正交基为基准，通过对数映射用黎曼法坐标表示数据集 X 上的所有点。由于黎曼法坐标的特点， p 到数据集 X 上其他点的距离和方向度量可得到很好的保持，这样就把高维空间的数据映射到低维空间。

算法步骤：

- 1) 从流形 M 的样本值 X 中选择一个基点 $p, p \in X$;
- 2) 选择 p 的 K 最近邻点表示为 $B(p)$ ，在 $B(p)$ 的外空间执行 PCA，得 $B(p)$ 在低维空间的黎曼法坐标及 p 点处切空间 T_pM 的一组标准正交基 $(p; e_1, e_2, \dots, e_d)$ ，即在最小二乘意义下寻找一组标准正交基 $(p; e_1, e_2, \dots, e_d)$ 使误差函数 F_{PCA} 最小如式(1)所示：

$$F_{\text{PCA}} = \sum_{i=1}^N \|x_i - \sum_{i=1}^d (x_i \cdot e_i) e_i\|^2 \quad (1)$$

3) 估计 $B(p)$ 上点到 X 上其他点的测地线距离。

定义一个加权无向图, 其数据点为顶点, 边缘连接相邻点, 边缘的权值为相邻点的欧式距离, 然后用 Dijkstra 算法来寻找最短路径, 估计测地线距离 $d(x, y)$ 其中 $(x, y) \in X \times B(p)$;

4) 利用上面步骤得到的值计算 p 到 $x \in X$ 中其他点的方向 g , 如式(2)所示:

$$g = \sum g^i e_i = \nabla_y d^2(x, y)|_{y=p} \quad (2)$$

5) 估计样本值 X 每个点的黎曼法坐标 Y , 如式(3)所示:

$$y_i = d(x, p) \frac{g^i}{|g|} \quad (3)$$

该算法是一种黎曼流形学习算法, 可以处理非零高斯曲率的数据, 适于三维人脸人脸识别。 $d=3$ 、 $k=14$ LOGMAP 对三维深度图降维如图 7 所示。

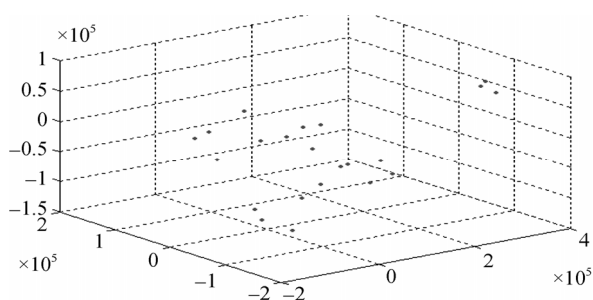


图 7 基于 LOGMAP 的三维深度图维数约减
Fig. 7 Dimension reduction of 3D depth image using LOGMAP

4 实验结果及分析

为了验证算法的有效性, 分别用 PCA、ISOMAP、LLE 和 LOGMAP 对 CIS 采集的三维深度图进行识别, 实验结果表明 LOGMAP 的识别结果优于其他 3 种学习方法。

三维深度数据集包含 7 个人的 49 幅多姿态(正/侧面, 抬头/低头)、多表情(睁/闭眼, 笑/不笑)、多细节(戴眼镜/不戴眼镜)以及多光照(亮光/暗光)的三维深度图。如图 8 所示。其训练集和测试集为: 每个人的前三幅图像为训练集, 训练集为 21, 后四幅图像为测试集, 测试集为 28。

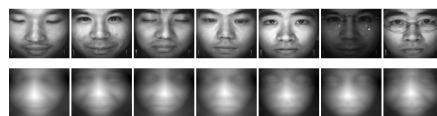


图 8 不同人脸的灰度图像和与之相对应的深度图像
Fig. 8 Different gray-scale face images and corresponding depth images

我们进行 2 组实验: 第一组是把高维空间的数据降到三维, 即 $d=3$ 时, k 取不同值 LOGMAP 进行识别, 维数约减结果如图 9, PCA、ISOMAP 和 LLE 分别降到三维作为对比实验, 识别率如表 1 所示; 第二组是把高维空间的数据降到二维, 即 $d=2$ 时, k 取不同值 LOGMAP 进行识别, 维数约减结果如图 10, PCA、ISOMAP 和 LLE 分别降到二维空间作为其对比实验其识别率如表 2 所示。维数约减图中的每一个点对应一幅图像, 基点(原点)对应待识别图像, 坐标轴仅表示点与点之间的相对距离, 即不同图像的相关性, 我们可以利用该相关性进行识别, 判别待识别的图是数据库中的哪幅图像。在这里 LOGMAP 仅取识别率前三的参数, 对比实验 LLE 和 ISOMA 仅取识别率最佳的参数。

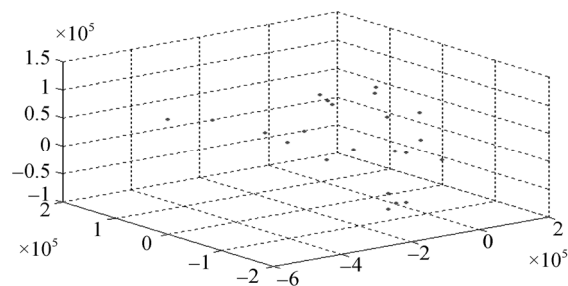


图 9 深度图的 LOGMAP 维数约减结果($k=17$)
Fig. 9 Dimension reduction result of LOGMAP with depth images($k=17$)

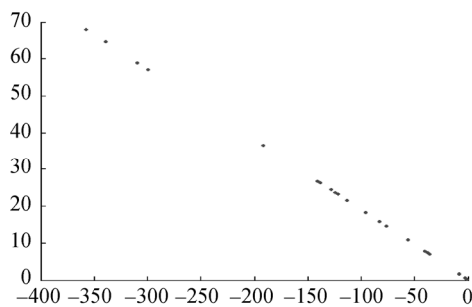


图 10 深度图像的 LOGMAP 维数约减结果($k=3$)
Fig. 10 Dimension reduction result of LOGMAP with depth images($k=3$)

表 1 深度图的识别率($d=3$)Table 1 Recognition rate of depth images ($d=3$)

算法	参数(k)	识别率/%
PCA		67.9
LLE	2	89.2
ISOMAP	2	96.4
	17	100
LOGMAP	3	96.4
	2	96.4

实验结果表明非线性黎曼流形学习算法 LOGMAP 在三维深度图识别比其他 3 种维数约减的识别率高, 适于三维深度图的识别, 但 LOGMAP 的识别率受参数影响较大, 不同的参数对识别率有较大的影响。

表 2 深度图的识别率($d=2$)Table 2 Rrecognition rate of depth image ($d=2$)

算法	参数(k)	识别率/%
PCA		50
LLE	3	89.2
ISOMAP	2	96.4
	3	96.4
LOGMAP	2	92.9
	12	85.7

5 结 论

三维人脸识别是人脸识别的热门研究领域, 流形学习是近年来提出的一类维数约减算法, 由于其具有非线性、直观的几何特性和收敛性等特点引起研究者的广泛关注, 基于流形学习的三维人脸识别可解决三维人脸识别的瓶颈问题, 意义重大。本文首先构建了基于流形学习的三维人脸识别框架, 并结合 LOGMAP 用于三维人脸深度图的识别, 实验结果表明 LOGMAP 在三维人脸深度图上可取的良好识别效果。

在本文的方法中也存在一些不足, 首先, LOGMAP 算法本身的距离和方向的求解不够精确, 有待改进; 其次, 在参数的设置还没有一个可靠的理论, 仅根据经验值来确定, 选择不同的参数对识别率有很大影响; 最后, 构建的流形学习框架仅适于人脸这样有相同感知的图像, 反之(例如字符)则不适用, 这些都是值得我们研究和思考的问题。

参考文献:

- [1] 王映辉. 人脸识别—原理、方法与技术[M]. 北京: 科学出版社, 2010: 290-292.
WANG Y H. Face recognition—principle, method and technique[M]. Beijing: Science Press, 2010: 290-292.
- [2] 叶长明, 蒋建国, 詹曙, 等. 基于三维成像系统的复数域人脸识别方法[J]. 电子测量与仪器学报, 2011, 25(5): 420-426.
YE CH M, JIANG J G, ZHAN SH, et al. Face recognition in complex domain based on 3D facial imaging system[J]. Journal of Electronic Measurement and Instrument, 2011, 25(5): 420-426.
- [3] JOLLIFFE I T. Principal component analysis[M]. Springer, 1989.
- [4] COX T, COX M. Multidimensional scanning[M]. Chapman and Hall, 1994.
- [5] TENENBAUM J, DE SILVA V, LANGFORD J. A global geometric framework for nonlinear dimensionality reduction[J]. Science, 2000: 290: 2319-2323.
- [6] LAW M H, JAIN A K. Incremental nonlinear dimensionality reduction by manifold learning[J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2006, 28(3): 377-391.
- [7] ROWEIS S T, SAUL L K. Nonlinear dimensionality reduction by locally linear embedding[J]. Science, 2000, 290(5500): 2323-2326.
- [8] 徐雪松, 张宏, 刘凤玉. 基于核函数距离测量的 LLE 降维及其在离散聚类中的应用[J]. 仪器仪表学报, 2008, 29(9): 1996-2000.
XU X S, ZHANG H, LIU F Y. LLE dimensionality reduction based on kernel-induced distance measurement and its application in clustering with outliers[J]. Chinese Journal of Scientific Instrument, 2008, 29(9): 1996-2000.
- [9] BELKIN M, NIYOGI P. Laplacian eigenmaps for dimensionality reduction and data representation[J]. Neural Computation, 2003, 15(6): 1373-1396.
- [10] ZHANG ZH Y, Zha Hongyuan. Principal manifolds and non linear dimensionality reduction via tangent space alignment[J]. SIAM Journal of Scientific Computing, 2004, 26(1): 313-338.
- [11] BRUN A, WESTIN C F, HERBERTHSON M, et al. Fast Manifold learning based on riemannian normal coordinates[C]. Proc. of 14th Scandinavian Conon Image Analysis, 2005: 921-929.
- [12] YE CH M, JIANG J G, et al. Face recognition with 3D real-time face imaging system based on LOGMAP[J]. 2011 10th International Conference on Electronic Measurement & Instruments (ICEMI), 2011: 373-376.
- [13] SEUNG S H, LEE D D. The Manifold of Perception[J].

Science, 2000: 2268-2269.

- [14] ZHANG J, LI S Z, WANG J. Manifold Learning and Application in Recognition[J]. Intelligent Multimedia Processing with Soft Computing, 2005, 168: 281-300.

- [15] 杜海顺, 柴秀丽, 汪凤泉, 等. 一种邻域保持判别嵌入人脸识别方法[J]. 仪器仪表学报, 2010, 31(3): 625-629.

DU H SH, CHAI X L, WANG F Q, et al. Face recognition method using neighborhood preserving discriminant embedding[J]. Chinese Journal of Scientific Instrument, 2010, 31(3): 625-629.

作者简介:

詹 曙: 1968 年出生, 博士, 现为合肥工业大学副教授。

主要研究方向为三维图像处理和分析、医学图像分析、模式识别。

E-mail: shu_zhan@hfut.edu.cn

Zhan Shu: born in 1968, Ph.D., associate professor in Hefei University of Technology. His research interests are 3D image processing and analysis, medical image analysis, pattern recognition.

张芝华: 1986 年出生, 毕业于安徽建筑工业学院, 现为合肥工业大学在读硕士研究生。主要研究方向为图像处理、流形学习及其理论应用。

E-mail: zhangzhihua789@126.com

Zhang Zhihua: born in 1986, graduated from Anhui University of Architecture. Now, he is a M.Sc. candidate in Hefei University of Technology. His main research direction is image processing, manifold learning and application.

安捷伦科技扩展任意波形发生器测试频段, 以增强其多功能性

该仪器提供宽带宽、高精度和多功能性, 可仿真各种真实的场景

安捷伦科技公司日前宣布为其业界领先的高分辨率、高带宽 M8190A 任意波形发生器添加增强功能。增强功能可为工程师提供更大的灵活性, 从而使用 5 到 7 GHz 频谱创建各种信号情景。

Agilent M8190A 是高保真度信号源, 能够同时提供宽带宽、高分辨率和高达 80 dBc 的无杂散动态范围。这种独一无二的组合使工程师可以创建各种信号情景, 提升设计性能, 增强分析能力。

AWG 是最全面的信号场景发生器, 提供在 8 GSa/s 14 位输出与 12 GSa/s 12 位输出之间轻松进行切换等诸多功能, 帮助工程师处理多种应用和满足不同的测量要求。

各种应用对信号特征的要求各不相同, 因此 Agilent M8190A 包含 3 个经优化的放大器, 以适用于不同应用。最新的增强功能进一步提高了该仪器的多功能性, 可支持各种输出格式, 包括 NRZ(不归零码)、DNRZ(双不归零码)、RZ(归零码)和双重模

式(doublet mode)。

从事各种应用并要求最高无杂散动态范围的工程师, 可以使用 DNRZ 格式。对于要求最佳脉冲性能的时域应用, 工程师可以使用 NRZ 格式。如果他们需要较高的频谱, 可切换到双重模式, 以便使用 5 到 7 GHz 频谱。从事射频应用的工程师, 能够在理想的频谱之间进行选择。

安捷伦数字与光学测试部总经理 Juergen Beck 表示: “M8190A 与安捷伦专利技术相结合, 可通过精密的信号仿真进行精确、可靠和可重复的测量。我们现在能够更进一步, 为客户提供更佳的功能。更高的信号保真度可以加快深入分析设计的速度, 并缩短测试周期。安捷伦致力于不断提供增强和创新功能, 为客户提供获得成功的工具。”

如欲了解关于安捷伦的详细信息, 请访问 www.agilent.com