

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2019 年 12 月 5 日 (05.12.2019)



(10) 国际公布号
WO 2019/227614 A1

- (51) 国际专利分类号:
G06K 9/00 (2006.01)
- (21) 国际申请号: PCT/CN2018/095462
- (22) 国际申请日: 2018 年 7 月 12 日 (12.07.2018)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201810557202.4 2018 年 6 月 1 日 (01.06.2018) CN
- (71) 申请人: 平安科技 (深圳) 有限公司 (PING AN TECHNOLOGY (SHENZHEN) CO., LTD.) [CN/CN];
中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。
- (72) 发明人: 赵峰 (ZHAO, Feng); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。
王健宗 (WANG, Jianzong); 中国广东省深圳市

福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。肖京 (XIAO, Jing); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

- (74) 代理人: 深圳市明日今典知识产权代理事务所 (普通合伙) (SHENZHEN MINGRIJINDIAN INTELLECTUAL PROPERTY AGENCY FIRM (GENERAL));
中国广东省深圳市南山区南头街道智恒新兴产业园 E 区 01B 栋 405 室, Guangdong 518000 (CN)。
- (81) 指定国 (除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,

(54) Title: METHOD AND DEVICE FOR OBTAINING TRIPLE OF SAMPLES, COMPUTER DEVICE AND STORAGE MEDIUM

(54) 发明名称: 样例三元组的获取方法、装置、计算机设备以及存储介质

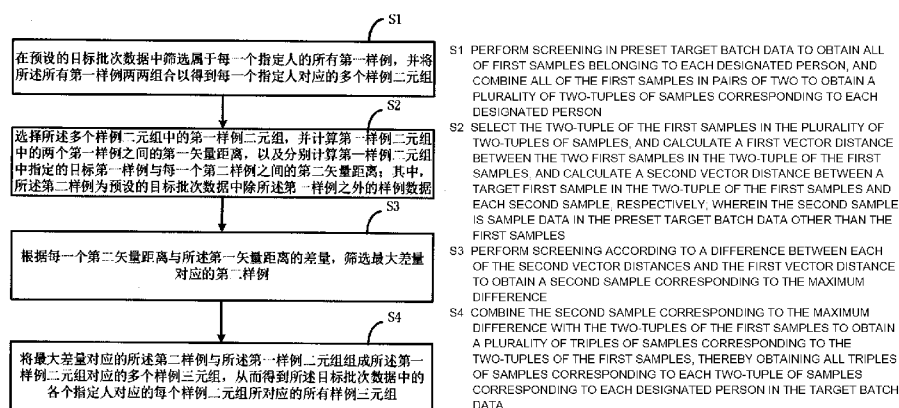


图 1

(57) Abstract: Disclosed in the present application is a method of obtaining a triple of samples, comprising: performing screening to obtain all of first samples of a designated person and combining the same in pairs of two to obtain a plurality of two-tuples of samples; selecting the two-tuple of the first samples, and calculating a first vector distance between the two first samples thereof and a second vector distance between a certain first sample and a second sample; performing screening according to a difference between the second vector distance and the first vector distance to obtain a second sample corresponding to the maximum difference, and combining the same with the two-tuple of the first samples to obtain a triple of samples.

(57) 摘要: 本申请揭示了样例三元组的获取方法, 包括: 筛选指定人的所有第一样例, 并两两组合得到多个样例二元组; 选择第一样例二元组, 并计算其两个第一样例的第一矢量距离, 以及某一第一样例与第二样例的第二矢量距离; 根据第二矢量距离与第一矢量距离的差量, 筛选最大差量对应的第二样例与第一样例二元组组成样例三元组。

SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, ZA, ZM, ZW。

- (84) 指定国 (除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,
IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告 (条约第21条(3))。

样例三元组的获取方法、装置、计算机设备以及存储介质

本申请要求于2018年6月01日提交中国专利局、申请号为2018105572024，发明名称为“样例三元组的获取方法、装置、计算机设备以及存储介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

5 技术领域

本申请涉及到机器学习领域，特别涉及到样例三元组的获取方法、装置、计算机设备以及存储介质。

背景技术

现有的样例三元组的选择是从整个训练数据中筛选的，导致样例三元组的总数量严重膨胀，其总数量的数量级跟样例数量的数量级是一个三次方的比例关系，导致现有方法获取出的样例三元组需要分批输入到网络模型中进行模型训练，但由于将每一分批的样例三元组送入训练模型中进行模型训练时，都会更改训练模型参数，使得训练模型训练时的各分批次的训练模型参量不同，使得后批次进入训练模型的样例三元组与训练模型的当前参量产生不匹配，样例三元组的效用逐渐下降。因此，从批次数量中合理获取样例三元组成为亟待解决的问题。

技术问题

15 本申请的主要目的为提供一种样例三元组的获取方法，旨在解决现有样例三元组的获取方法导致样例三元组的数量过于庞大影响训练模型的训练过程。

技术解决方案

本申请提出一种样例三元组的获取方法，包括：

20 在预设的目标批次数据中筛选属于每一个指定人的所有第一样例，并将所述所有第一样例两两组合以得到每一个指定人对应的多个样例二元组；

选择所述多个样例二元组中的第一样例二元组，并计算第一样例二元组中的两个第一样例之间的第一矢量距离，以及分别计算第一样例二元组中指定的目标第一样例与每一个第二样例之间的第二矢量距离；其中，所述第二样例为预设的目标批次数据中除所述第一样例之外的样例数据；

根据每一个第二矢量距离与所述第一矢量距离的差量，筛选最大差量对应的第二样例；

25 将最大差量对应的第二样例与所述第一样例二元组组成所述第一样例二元组对应的多个样例三元组，从而得到所述目标批次数据中的各个指定人对应的每个样例二元组所对应的所有样例三元组。

本申请还提供了一种样例三元组的获取装置，包括：

第一筛选模块，用于在预设的目标批次数据中筛选属于每一个指定人的所有第一样例，并将所述所有第一样例两两组合以得到每一个指定人对应的多个样例二元组；

30 计算模块，用于选择所述多个样例二元组中的第一样例二元组，并计算第一样例二元组中的两个第一样例之间的第一矢量距离，以及分别计算第一样例二元组中指定的目标第一样例与每一个第二样例之

间的第二矢量距离；其中，所述第二样例为预设的目标批次数据中除所述第一样例之外的样例数据；

第二筛选模块，用于根据每一个第二矢量距离与所述第一矢量距离的距离差，筛选最大距离差对应的第二样例；

第一组成模块，用于将最大距离差对应的第二样例与样例二元组组成第一样例二元组对应的多个样例三元组，从而得到目标批次数据中的各个指定人对应的每个样例二元组所对应的所有样例三元组。

本申请还提供了一种计算机设备，包括存储器和处理器，所述存储器存储有计算机可读指令，所述处理器执行所述计算机可读指令时实现上述方法的步骤。

本申请还提供了一种计算机非易失性可读存储介质，其上存储有计算机可读指令，所述计算机可读指令被处理器执行时实现上述的方法的步骤。

10 有益效果

本申请通过将几万或几十万数量级的整个训练数据分成包含指定数量样例的小批次数据，然后在小批次数据中获取样例三元组，避免了样例三元组数量过于庞大而导致的模型训练速度缓慢、甚至无法顺畅训练的情况，加快训练模型的训练速度，且同时限定只获取与AP对矢量距离最大的候选N组成APN，进一步减少获取样例三元组的数量，提高训练速度，同时基于与AP对矢量距离最大的候选N组成的样例三元组APN的具有更好的数据代表性，在进一步提升训练速度的同时，优化训练模型的效果，使得筛选的样例三元组在整个训练数据中具有代表性，且对训练模型的训练过程更优，表现为训练模型的参量一致性好；本申请在较小的批次数据范围内获取样例三元组，所产生的样例三元组可作为小批量随机梯度下降方法进行深度学习模型训练时的一个批量数据，一次性全部送入训练模型进行训练，避免了多次输入时训练模型参量发生改变，导致样例三元组相对于当前参量下的训练模型效用下降的情况；本申请在较小批次数据范围内获取样例三元组且只选择与AP对矢量距离最大的候选N组成APN，使得获取的样例三元组和随后使用样例三元组进行模型训练可以连续进行，避免了因为训练数据太多，导致获取样例三元组后需要额外花费计算时间对样例三元组进行分批后再送入训练模型进行训练，提高了训练速度。

附图说明

- 图1 本申请一实施例的样例三元组的获取方法的流程示意图；
- 图2 本申请一实施例的样例三元组的获取装置的结构示意图；
- 图3 本申请一实施例的第一筛选模块的结构示意图；
- 图4 本申请另一实施例的样例三元组的获取装置的结构示意图；
- 图5 本申请一实施例的输入模块的结构示意图；
- 图6 本申请一实施例的第一输入单元的结构示意图；
- 图7 本申请另一实施例的输入模块的结构示意图；
- 图8 本申请另一实施例的第二输入单元的结构示意图；

图9 本申请一实施例的计算机设备的内部结构示意图。

本发明的最佳实施方式

参照图 1, 本申请一实施例的样例三元组的获取方法, 包括:

S1: 在预设的目标批次数据中筛选属于每一个指定人的所有第一样例, 并将所述所有第一样例两两组合以得到每一个指定人对应的多个样例二元组。

本实施例的指定人为泛指, 上述指定人的样例数据为批次数据的子集, 本实施例通过定义指定人的样例数据为第一样例, 以区别于批次数据中指定人之外的其他人的样例数据, 此处“第一”仅为区别, 不作限定。举例地, 批次数据 A 中, 包括 a、b、c、d、e、f 等 M 个人的样例数据, 上述指定人可以为 a、b、c、d、e、f 等人中的任意一个, 比如指定 a 为指定人, 则将属于 a 的所有样例数据按照两两组合的方式组配二元组 AP 对 (Anchor, Positive 对), 比如, 属于 a 的所有样例数据为 10 个, 则 AP 对的数量对应为 C_2^{10} ; 除 a 之外的其他人的样例数据成为样例三元组 APN 的候选 N (Negative)。本实施例通过分别对 M 个人中分别属于各自的样例数据, 分别进行两两组合形成每个人相对的二元组 AP 对, 然后再筛选合适的 N 组成样例三元组 APN, 以便形成可用于训练模型训练的、具有数据代表性的样例三元组数据集, 本实施例通过在批次数据中取最少量、最具代表性的样例三元组, 提高训练模型的训练速度及训练效果, 提高训练模型的训练时效性以及实用性。

S2: 选择所述多个样例二元组中的第一样例二元组, 并计算第一样例二元组中的两个第一样例之间的第一矢量距离, 以及分别计算第一样例二元组中指定的目标第一样例与每一个第二样例之间的第二矢量距离; 其中, 所述第二样例为预设的目标批次数据中除所述第一样例之外的样例数据。

本实施例通过定义批次数据中指定人之外的样例数据为第二样例, 以区别于第一样例, 此处“第二”同上述“第一”仅为区别, 不作限定其他部分的“第一”、“第二”的作用相同, 不赘述。本实施例的第一矢量距离为第一二元组 AP 对的矢量距离, 第二矢量距离为第一样例二元组 AP 对中的 A 与候选 N 之间的矢量距离, 其中 A 为第一样例二元组中指定的第一样例。

S3: 根据每一个第二矢量距离与所述第一矢量距离的差量, 筛选最大差量对应的第二样例。

本实施例对候选 N 的筛选条件限定为第二矢量距离 AN 与第一矢量距离 AP 相距距离最大的候选 N, 即通过 $\max(s_i^{ap} - s_i^{an} + a, 0)$ 筛选候选 N, 其中, a 是常量, s_i^{ap} 为 x_a 和 x_p 之间的欧式距离, s_i^{an} 为 x_a 和 x_n 之间的欧式距离; s_i^{ap} 计算公式为 $s_i^{ap} = \sqrt{(x_{a1} - x_{p1})^2 + (x_{a2} - x_{p2})^2 + \dots + (x_{an} - x_{pn})^2}$; x_a 为 Anchor 对应的矢量, x_p 为 Positive 对应的矢量; s_i^{an} 计算过程同上, 将上述公式中的 x_p 替换为 x_n 即可, 其中 x_n 为 Negative 对应的矢量。本实施例的候选 N 不仅满足第二矢量距离与所述第一矢量距离的距离差大于零, 且在大于零的前提下, 进一步优选第二矢量距离与所述第一矢量距离的差量最大的情况下对应的唯一候选 N, 以进一步减少样例三元组的数量, 一个样例二元组对应形成一个样例三元组, 进一步提高筛选候选 N 后组成的样例三元组 APN 的数据代表性, 且进一步减少样例三元组 APN 的总数量, 以提

高模型训练中的训练速度和训练效果。第二矢量距离与所述第一矢量距离的差量最大，以进一步限定候选 N 的获取范围，只选差量最大时对应的唯一候选 N，确保形成的样例三元组的总数量最小，最大化消除了样例三元组数量过分膨胀而影响模型训练的问题，且同时确保组配的样例三元组 APN 具有代表性，能够充分的代表训练数据数据中的全部数据分布状态，保障训练模型的有效性。本实施例优选第二矢量

5 距离与所述第一矢量距离的差量大于零的情况下，且选择差量最大的候选 N，大量缩减候选 N 的数量，进一步减少训练模型的数据输入量，提高模型训练速度；候选 N 与 A 或 P 相距最大距离，以最大化减少数据量较少情况下的数据偏差，提高筛选后的最少量的样例三元组集合在整个批次数据中的代表性，即能够充分的代表训练数据中的全部数据分布状态，确保通过较少量样例三元组训练模型中获得的训练模型的权重有效性。

10 S4：将最大差量对应的所述第二样例与所述第一样例二元组组成第一样例二元组对应的多个样例三元组，从而得到所述目标批次数据中的各个指定人对应的每个样例二元组所对应的所有样例三元组。

本实施例通过在批次数据中筛选合适的样例三元组作为训练模型的训练数据输入，大大减少了训练模型在训练过程中需要输入的数据量，加快了训练模型的训练速度，避免了将所有训练数据输入到训练模型中进行训练时易发生的训练缓慢或训练堵塞中断等现象。而且本申请通过只获取与 AP 对矢量距离

15 最大的候选 N 组成 APN，最大化减少获取样例三元组的数量，并同时优化样例三元组的组配合理性，减少训练数据偏差，提高样例三元组在整个训练数据中的代表性，使得输入较少的训练数据量却训练得到更具有实用性的训练模型。本实施例对批次数据中的任一指定人的样例数据进行排列组合形成 AP 对，并将指定人的任一 AP 对分别与 AP 对矢量距离最大的候选 N 组成样例三元组 APN，通过对批次数据中所有人的所有 AP 对进行组配样例三元组 APN，将批次数据优化成具有一定数量的样例三元组 APN 的集合，

20 以进一步优化输入到训练模型中的训练数据组成与分布，提高训练模型的训练效果。

进一步地，上述步骤 S1，包括：

S10：将训练数据拆分成包含指定数量样例的多个批次数据。

本步骤的训练数据是指常规用于深度学习训练模型的训练数据的所有数据，训练数据总量包括几万或几十万数量级。本实施例通过将几万或几十万数量级的大数量级的训练数据拆分成多个小数量的批次数据，以进一步限制组合的样例三元组的数量，以提高训练模型的训练效率，节省训练模型的训练耗费时间。本实施例的指定数量兼顾了训练模型在训练时的硬件设备计算能力和训练模型性能要求的两方面因素，本实施例优选指定数量为不大于 128 个样例数据，使得样例三元组只能在 128 个样例数据的批次数据范围中进行获取，经过排列组合形成的样例三元组的数量得到有效控制，使得训练模型在训练过程中能在 GPU 的有限储存容量下，可保证持续有效的高速计算能力，加快训练模型的训练速度；同时由于批

25 次数据的总数量有限，使得每个批次数据中的样例三元组的数量不会过分膨胀，使得每个批次数据中产生的样例三元组集合中的所有样例三元组，都能一次性输入到训练模型中进行模型训练，避免了大批次

30

训练数据分多批次输入时，会导致每批次输入数据时发生的训练模型参数跟随变更，使得样例三元组对训练模型的训练效用下降的弊端，本实施例通过将大数量级的数据先分成各批次数据，批次数据中的样例数据再形成样例三元组，有效控制样例三元组的总量，使得样例三元组可一次性输入到训练模型中进行模型训练，保证了训练模型的参数稳定性，使得训练模型具有良好的应用效果，表现为等错误率最低

5 （等错率即正确数据的错误辨识误差与错误数据的误判通过误差相等）。本实施例的批次数据中包括 M 个人的样例数据，每个人的样例数据为 R 个， M 取值远小于训练数据中的总人数，比如 M 为 30 至 60 之间， R 为 20-40 之间，则满足批次数据的全部样例数量不大于 128，其中， R 包括相同的样例数据和不同的样例数据，当 R 个中的样例数据均不同时更能优化样例数据的分布状态，提高训练模型的优化效果；但若某一人的不同样例数据小于 R ，为满足数量 R 的要求，可允许使用部分重复数据。

10 S11：将所述多个批次数据中的目标批次数据输入到预设训练模型中，通过预设训练模型的正向训练获取到所述每个批次数据中所述指定数量样例分别对应的矢量。

本实施例通过训练模型的正向训练计算各输入样例分别对应的矢量值，以便根据两个样例之间的矢量距离合理组合样例三元组，提高样例三元组的组配合理性，减小各批次数据之间的偏差，提高训练模型的训练效果，比如：提高二分类训练模型的判断准确率。

15 S12：筛选所述目标批次数据中属于某一指定人的所有第一样例，并将所述所有第一样例进行两两组合成多个样例二元组，其中第一样例包含与其对应的矢量。

本实施例通过将某一指定人的所有样例数据定义为第一样例，并在第一样例中指定一个为 A ，其他的第一样例分别为 P ，依次排列组合形成 AP 对，并根据第一样例分别对应的各矢量，获取各 AP 对对应的第一矢量距离 AP ，根据第一矢量距离 AP 在批次数据中筛选差量最大的候选 N ，组成样例三元组 APN 。

20 进一步地，上述步骤 S4 之后，包括：

S41：将训练数据中的多个批次数据中所有第一样例二元组分别对应的所有样例三元组汇总成样例三元组训练集。

25 本实施例先获取第一样例二元组与其对应的满足预设条件的第二样例，组成第一样例二元组对应的样例三元组集合。本实施例通过训练数据中的多个所述批次数据分别对应的第一子集汇总成样例三元组训练集，以获取从整个训练数据中优化筛选的样例三元组数据作为训练集。

S42：将所述样例三元组训练集输入以 Triplet Loss 为损失函数的二分类网络模型中进行模型训练。

30 本实施例的二分类网络模型是指只对输入数据进行‘是或否’判断的网络模型，举例地，人脸识别模型或说话人确认模型均为上述的二分类网络模型。本实施例的 Triplet Loss 具有三个参数，则训练过程中需要同时输入三个样例数据才能完成训练模型的训练过程，因此本实施例的样例三元组 APN 的挑选方式对以 Triplet Loss 为损失函数的网络优化效果更加显著。本实施例将样例数据表示为 (s_i, l_i) ， $i=1, 2, 3, \dots, N$ ，其中 s_i 是样例数据本身， l_i 是其对应的待测人。本实施例的 Triplet Loss 其具体计算方

法如下：将所有样例数据 (s_i, l_i) 都输入到二分类网络模型中，得到各样例数据分别对应输出的向量 x_i ；从上述向量 x_i 中随机选取一个向量作为 x_a ，然后再随机选取另一个向量作为 x_p ， x_a 和 x_p 应该属于同一个人；计算 x_a 和 x_p 之间的欧式距离 s_i^{ap} ；设向量 $x_i = x_{i1}, x_{i2}, \dots, x_{in}$ ，则 s_i^{ap} 计算公式为 $s_i^{ap} = \sqrt{(x_{a1} - x_{p1})^2 + (x_{a2} - x_{p2})^2 + \dots + (x_{an} - x_{pn})^2}$ ；随机选取另一个向量作为 x_n ， x_a 和 x_n 分别

5 属于不同的人；计算 x_a 和 x_n 之间的欧式距离 s_i^{an} ；则损失函数 L 由下面的公式计算：

$$L = \sum_{i=0}^N \max(s_i^{ap} - s_i^{an} + a, 0); a \text{ 是常量取值范围在 } 0.1 \text{ 至 } 0.25 \text{ 之间, } N \text{ 是选取的样例三元组的个数。}$$

进一步地，所述第一样例和所述第二样例为说话人的话语样例，所述二分类网络模型为说话人确认网络模型；上述步骤 S42，包括：

10 S421：将说话人的话语样例组成的样例三元组训练集输入以 Triplet Loss 为损失函数的说话人确认网络模型中进行模型训练。

本实施例以用于说话人确认为例，人说话时发音器官和发音习惯的差异都以复杂的形势反映在说话人的语音特征中，使得每个人的话语都带有强烈的个人色彩，本实施例通过将话语样例中的语义信息平均化，挖掘出包含在话语样例中的说话人的个性因素，强调不同人之间的语音特征的特征差异。批次数据中 M 为 40 个人， R 为 30 个不同语音样例，比如正常说话状态下选择 10 个不同语音样例，愤怒状态下选择 10 个不同语音样例，伤心状态下选择 10 个不同语音样例，将指定人的语音样例与其他人的语音样例组成的样例三元组，指定数量为 128，说话人的样例三元组包括说话人的话语样例组成的样例三元组。通过训练说话人确认网络模型训练，形成提取话语样例中的语速、音质、语音波形等语音特征的提取方式，通过比较语音特征的相似度，对待测说话人与预注册说话人进行一对一的确认，输出待测说话人与预注册说话人是同一人或待测说话人与预注册说话人不是同一人的确认结果。本实施例通过上述方法在

20 众多说话人的话语样例中，先分成包含指定数量说话人的批次数据，并在批次数据中筛选合适的样例三元组 APN，提高说话人确认网络模型的训练速度以及训练效果，以便扩展说话人确认网络模型的应用领域，包括在金融、证券、社保、公安、军队及其他民用安全认证等行业和部门满足相应的使用需求。

进一步地，上述步骤 S421，包括：

S4211：将待测人的话语样例输入到训练好的所述说话人确认网络模型中。

25 本实施例的待测人的话语样例包括待测人的自然语音、合成语音等包含音质等一种或多种语音特征的语音样例。举例地，合成语音包括经过去除噪音处理的话语样例、经过去噪音处理并二次排列拼接的话语样例以及去除噪音后添加了新的背景音等。合成语音中带有强烈的个人色彩的语音特征中至少一种语音特征保持不变，则可实现说话人确认。

30 S4212：在说话人确认网络模型中计算待测人的话语样例中的第一语音特性与预注册人的第二语音特性的第一相似度。

本实施例的待测人的话语样例中的第一语音特性与预注册人的第二语音特性的第一相似度，可通过欧式距离、马氏距离等进行判定，优选通过欧式距离进行判定。本实施例通过 $\max(s_i^{ap} - s_i^{an} + a, 0)$ 筛选候选 N，其中，a 是常量， s_i^{ap} 为 x_a 和 x_p 之间的欧式距离， s_i^{an} 为 x_a 和 x_n 之间的欧式距离； x_a 为话语样例批次数据中的 Anchor 对应的矢量， x_p 为 Positive 对应的矢量； x_n 为 Negative 对应的矢量。通过计算第二矢量距离与所述第一矢量距离的距离差，在距离差大于零的候选 N 集合中，优选距离差最大的候选 N，即优选与指定人的语音特性差别最大的人的语音样例，以进一步提高的样例三元组 APN 的数据代表性，且减少样例三元组 APN 的总数量，提高说话人确认网络模型训练中的训练速度和训练效果。

S4213：判断所述第一相似度是否在预设阈值范围内。

举例地，预设范围为待测人的话语样例中的第一语音特性与预注册人的第二语音特性的欧式距离大于 0.9，则第一相似度高或判定第一语音特性的待测人与第二语音特性对应定的预注册人为同一人。

S4214：若是，则判定待测人与预注册人为同一个人。

本申请另一实施例中，所述第一样例和所述第二样例为人脸影像的样例，所述二分类网络模型为人脸识别网络模型；上述步骤 S421，包括：

S4215：将人脸影像组成的样例三元组训练集输入以 Triplet Loss 为损失函数的人脸识别网络模型中进行模型训练。

本实施例以用于人脸识别为例，样例为人脸影像，用于训练人脸识别模型，并根据人脸识别模型实现对指定待识别人的人脸影像的一对一判定，即通过便是结果为是或否，以便确定待识别人与注册人是否为同一个人。本实施例的批次数据中 M 为 50 个人，R 为 20 个不同角度的人脸影像，比如，正面人脸影像旋转到正右侧脸影像之间选择 10 个不同人脸影像，正面人脸影像旋转到正左侧脸影像之间再选择 10 个不同人脸影像。将指定的人脸影像与其他的人脸影像组成的样例三元组。通过对训练人脸识别网络模型训练，形成提取人脸影像中的人脸特征的提取方式，以便通过比较人脸特征的相似度，对待测人与预注册人进行一对一的确认，输出待测人与预注册人是同一人或待测人与预注册人不是同一人的确认结果。本实施例通过上述方法在 50 个人的人脸影像组成的批次数据中筛选合适的样例三元组 APN，提高人脸识别网络模型的训练速度以及训练效果，以便扩展人脸识别网络模型的应用领域，以便在金融、证券、社保、公安、军队及其他民用安全认证等行业和部门满足使用需求。

进一步地，上述步骤 S4215 之后，包括：

S4216：将待测人的人脸影像输入到训练好的所述人脸识别网络模型中。

本实施例的待测人的人脸影像包括待测人脸五官的五官特征，比如五官各自的大小、姿态、位置等。

S4217：在人脸识别网络模型中计算待测人的人脸影像中的第一人脸特征与预注册人的第二人脸特征的第二相似度。

本实施例通过 $\max(s_i^{ap} - s_i^{an} + a, 0)$ 筛选候选 N，其中，a 是常量， s_i^{ap} 为 x_a 和 x_p 之间的欧式距离，

s_i^{an} 为 x_a 和 x_n 之间的欧式距离； x_a 为人脸影像批次数据中的 Anchor 对应的矢量， x_p 为 Positive 对应的矢量； x_n 为 Negative 对应的矢量。通过计算第二矢量距离与所述第一矢量距离的距离差，在距离差大于零的候选 N 集合中，进一步优选距离差最大的候选 N，即优选与指定人的人脸特性差别最大的人的人脸影像，以进一步提高的样例三元组 APN 的数据代表性，且减少样例三元组 APN 的总数量，以提高人脸

5 识别网络模型训练中的训练速度和训练效果。

S4218：判断所述第二相似度是否在预设阈值范围内。

本实施例的第二相似度可通过欧式距离、马氏距离等进行判定，优选通过欧式距离进行判定。举例地，本实施例的预设范围为 0.8 或以上，欧氏距离越大表示相似度高或判定待测人与预注册人为同一人。

S4219：若是，则判定待测人与预注册人为同一个人。

10 进一步地，上述步骤 S4214 或步骤 S4219 之后，包括：

S4220：控制装配所述说话人确认网络模型的安全认证系统展示通过认证。

本实施例通过将说话人确认网络模型的确认结果，转换成与其相关联的安全认证系统的控制指令，以提高说话人确认系统的实际应用效果。举例地，说话人确认网络模型的确认待测人与预注册人为同一个人后，待测人可打开安全认证系统控制的应用平台，使应用平台处于可使用的状态。本实施例展示的
15 通过认证的方式包括对话框文本显示和/或语音播报和/或警示铃鸣声和/或警示灯的闪光方式等。

本实施例本申请通过将几万或几十万数量级的整个训练数据分成包含指定数量样例的小批次数据，然后在小批次数据中获取样例三元组，避免了样例三元组数量过于庞大而导致的模型训练速度缓慢、甚至无法顺畅训练的情况，加快训练模型的训练速度，且同时限定只获取与 AP 对矢量距离最大的候选 N 组成 APN，进一步减少获取样例三元组的数量，提高训练速度，同时基于与 AP 对矢量距离最大的候选 N 组成的样例三元组 APN 的具有更好的数据代表性，在进一步提升训练速度的同时，优化训练模型的效果，使得筛选的样例三元组在整个训练数据中具有代表性，且对训练模型的训练过程更优，表现为训练模型的参量一致性好；通过在较小的批次数据范围内获取样例三元组，所产生的样例三元组可作为小批量随机梯度下降方法进行深度学习模型训练时的一个批量数据，一次性全部送入训练模型进行训练，避免了多次输入时训练模型参量发生改变，导致样例三元组相对于当前参量下的训练模型效用下降的情况；在较
20 小批次数据范围内获取样例三元组且只选择与 AP 对矢量距离最大的候选 N 组成 APN，使得获取的样例三元组和随后使用样例三元组进行模型训练可以连续进行，避免了因为训练数据太多，导致获取样例三元组后需要额外花费计算时间对样例三元组进行分批后再送入训练模型进行训练，提高了训练速度。

参照图 2，本申请一实施例的样例三元组的获取装置，包括：

第一筛选模块 1，用于在预设的目标批次数据中筛选属于每一个指定人的所有第一样例，并将所述
30 所有第一样例两两组合以得到每一个指定人对应的多个多个样例二元组。

本实施例的指定人为泛指，上述指定人的样例数据为批次数据的子集，通过定义指定人的样例数据

为第一样例，区别于批次数据中指定人之外的其他人的样例数据。举例地，批次数据 A 中，包括 a、b、c、d、e、f 等 M 个人的样例数据，上述指定人可以为 a、b、c、d、e、f 等人中的任意一个，比如指定 a 为指定人，则将属于 a 的所有样例数据按照两两组合的方式组配二元组 AP 对 (Anchor, Positive 对)，比如，属于 a 的所有样例数据为 10 个，则 AP 对的数量对应为 C_7^2 ；除 a 之外的其他人的样例数据成为样
 5 例三元组 APN 的候选 N (Negative)。本实施例通过分别对 M 个人中分别属于各自的样例数据，分别进行两两组合形成每个人相对的二元组 AP 对，然后再筛选合适的 N 组成样例三元组 APN，以便形成可用于训练模型训练的、具有数据代表性的样例三元组数据集合，提高训练模型的训练时效性以及实用性。

计算模块 2，用于选择所述多个样例二元组中的第一样例二元组，并计算第一样例二元组中的两个第一样例之间的第一矢量距离，以及分别计算第一样例二元组中指定的目标第一样例与每一个第二样例
 10 之间的第二矢量距离；其中，所述第二样例为预设的目标批次数据中除所述第一样例之外的样例数据。

本实施例通过定义批次数据中某一指定人之外的样例数据为第二样例，以区别于第一样例，此处“第二”同上述“第一”仅为区别，不作限定其他部分的“第一”、“第二”的作用相同，不赘述。本实施例的第一矢量距离为第一二元组 AP 对的矢量距离，第二矢量距离为第一样例二元组 AP 对中的 A 与候选
 N 之间的矢量距离，其中 A 为第一样例二元组中指定的第一样例。

15 第二筛选模块 3，用于根据每一个第二矢量距离与第一矢量距离的差量，筛选最大差量对应的第二样例。

本实施例对候选 N 的筛选条件限定为第二矢量距离 AN 与第一矢量距离 AP 相距距离最大的候选 N，即通过 $\max(s_i^{ap} - s_i^{an} + a, 0)$ 筛选候选 N，其中，a 是常量， s_i^{ap} 为 x_a 和 x_p 之间的欧式距离， s_i^{an} 为 x_a 和 x_n 之间的欧式距离； s_i^{ap} 计算公式为 $s_i^{ap} = \sqrt{(x_{a1} - x_{p1})^2 + (x_{a2} - x_{p2})^2 + \dots + (x_{an} - x_{pn})^2}$ ； x_a 为
 20 Anchor 对应的矢量， x_p 为 Positive 对应的矢量； s_i^{an} 计算过程同上，将上述公式中的 x_p 替换为 x_n 即可，其中 x_n 为 Negative 对应的矢量。本实施例的候选 N 不仅满足第二矢量距离与所述第一矢量距离的距离差大于零，且在大于零的前提下，优选第二矢量距离与所述第一矢量距离的差量最大的情况下对应的唯一候选 N，以进一步减少样例三元组的数量，一个样例二元组对应形成一个样例三元组，进一步提高筛选候选 N 后组成的样例三元组 APN 的数据代表性，且减少样例三元组 APN 的总数量，以提高模型训练中的
 25 的训练速度和训练效果。第二矢量距离与所述第一矢量距离的差量最大，以限定候选 N 的获取范围，只选差量最大时对应的唯一候选 N，确保形成的样例三元组的总数量最小，最大化消除了样例三元组数量过分膨胀而影响模型训练的问题，且同时确保组配的样例三元组 APN 具有代表性，能够充分的代表训练数据数据中的全部数据分布状态，保障训练模型的有效性。本实施例优选第二矢量距离与所述第一矢量距离的差量大于零的情况下，且选择差量最大的候选 N，大量缩减候选 N 的数量，减少训练模型的数据
 30 输入量，提高模型训练速度；候选 N 与 A 或 P 相距最大距离，以最大化减少数据量较少情况下的数据偏差，提高筛选后的最少量的样例三元组集合在整个批次数据中的代表性，即能够充分的代表训练数据中

的全部数据分布状态，通过较少量样例三元组训练模型中获得的训练模型的权重有效性。

第一组成模块 4，用于将最大差量对应的第二样例与第一样例二元组组成第一样例二元组对应的多个样例三元组，得到目标批次数据中的各个指定人对应的每个样例二元组所对应的所有样例三元组。

本实施例通过在批次数据中筛选合适的样例三元组作为训练模型的训练数据输入，大大减少了训练模型在训练过程中需要输入的数据量，加快了训练模型的训练速度，避免了将所有训练数据输入到训练模型中进行训练时易发生的训练缓慢或训练堵塞中断等现象。而且本申请通过只获取与 AP 对矢量距离最大的候选 N 组成 APN，最大化减少获取样例三元组的数量，并同时优化样例三元组的组配合理性，减少训练数据偏差，提高样例三元组在整个训练数据中的代表性，使得输入较少的训练数据量却训练得到更具有实用性的训练模型。本实施例对批次数据中的任一指定人的样例数据进行排列组合形成 AP 对，并将指定人的任一 AP 对分别与 AP 对矢量距离最大的候选 N 组成样例三元组 APN，通过对批次数据中所有人的所有 AP 对进行组配样例三元组 APN，将批次数据优化成具有一定数量的样例三元组 APN 的集合，以进一步优化输入到训练模型中的训练数据组成与分布，提高训练模型的训练效果。

参照图 3，上述第一筛选模块 1，包括：

拆分单元 10，用于将训练数据拆分成包含指定数量样例的多个批次数据。

本实施例的训练数据是指常规用于深度学习训练模型的训练数据的所有数据，训练数据总量包括几万或几十万数量级。本实施例通过将几万或几十万数量级的大数量级的训练数据拆分成多个小数量的批次数据，以进一步限制组合的样例三元组的数量，以提高训练模型的训练效率。本实施例的指定数量兼顾了训练模型在训练时的硬件设备计算能力和训练模型性能要求的双方面因素，本实施例优选指定数量为不大于 128 个样例数据，使得样例三元组只能在 128 个样例数据的批次数据范围中进行获取，经过排列组合形成的样例三元组的数量得到有效控制，使得训练模型在训练过程中能在 GPU 的有限储存容量下，可保证持续有效的高速计算能力，加快训练模型的训练速度；同时由于批次数据的总数量有限，使得每个批次数据中的样例三元组的数量不会过分膨胀，使得每个批次数据中产生的样例三元组集合中的所有样例三元组，都能一次性输入到训练模型中进行模型训练，避免了大批次训练数据分多批次输入时，会导致每批次输入数据时发生的训练模型参数跟随变更，使得样例三元组对训练模型的训练效用下降的弊端，本实施例通过将大数量级的数据先分成各批次数据，批次数据中的样例数据再形成样例三元组，有效控制样例三元组的总量，使得样例三元组可一次性输入到训练模型中进行模型训练，保证了训练模型的参数稳定性，使得训练模型具有良好的应用效果，表现为等错误率最低。本实施例的批次数据中包括 M 个人的样例数据，每个人的样例数据为 R 个，M 取值远小于训练数据中的总人数，比如 M 为 30 至 60 之间，R 为 20-40 之间，则满足批次数据的全部样例数量不大于 128，其中，R 包括相同的样例数据和不同的样例数据，当 R 个中的样例数据均不同时更能优化样例数据的分布状态，提高训练模型的优化效果；但若某一人的不同样例数据小于 R，为满足数量 R 的要求，可允许使用部分重复数据。

获取单元 11, 用于将所述多个批次数据中的目标所述批次数据输入到预设训练模型中, 通过预设训练模型的正向训练获取到所述每个批次数据中所述指定数量样例分别对应的矢量。

本实施例通过训练模型的正向训练计算各输入样例分别对应的矢量值, 以便根据两个样例之间的矢量距离合理组合样例三元组, 提高样例三元组的组配合理性, 减小各批次数据之间的偏差, 提高训练模型 5 的训练效果, 比如: 提高二分类训练模型的判断准确率。

筛选单元 12, 用于筛选所述目标批次数据中属于某一指定人的所有第一样例, 并将所述所有第一样例进行两两组合成多个样例二元组, 其中第一样例包含与其对应的矢量。

本实施例通过将某一指定人的所有样例数据定义为第一样例, 并在第一样例中指定一个为 A, 其他的第一样例分别为 P, 依次排列组合形成 AP 对, 并根据第一样例分别对应的各矢量, 获取各 AP 对对应的第一矢量距离 AP, 以便根据第一矢量距离 AP, 在批次数据中筛选差量最大时对应的候选 N, 组成样例三元组 APN。

参照图 4, 本申请另一实施例的样例三元组的获取装置, 包括:

汇总模块 41, 用于将训练数据中的多个批次数据分别对应的第一子集汇总成样例三元组训练集。

本实施例先获取第一样例二元组与其对应的满足预设条件的第二样例, 组成第一样例二元组对应的 15 样例三元组集合。本实施例通过训练数据中的多个所述批次数据分别对应的第一子集汇总成样例三元组训练集, 以获取从整个训练数据中优化筛选的样例三元组数据作为训练集。

输入模块 42, 用于将所述样例三元组训练集输入以 Triplet Loss 为损失函数的二分类网络模型中进行模型训练。

本实施例的二分类网络模型是指只对输入数据进行‘是或否’判断的网络模型, 举例地, 人脸识别 20 模型或说话人确认模型均为上述的二分类网络模型。本实施例的 Triplet Loss 具有三个参数, 则训练中需要同时输入三个样例数据, 完成训练模型的训练过程, 因此本实施例的样例三元组 APN 的挑选方式对以 Triplet Loss 为损失函数的网络优化效果更加显著。本实施例将样例数据表示为 (s_i, l_i) , $i=1, 2, 3, \dots, N$, 其中 s_i 是样例数据本身, l_i 是其对应待测人。本实施例的 Triplet Loss 其具体计算方法如下: 将所有样例数据 (s_i, l_i) 都输入到二分类网络模型中, 得到各样例数据分别对应输出的矢量 x_i ; 25 从上述矢量 x_i 中随机选取一个矢量作为 x_a , 然后再随机选取另一个矢量作为 x_p , x_a 和 x_p 应该属于同一个人; 计算 x_a 和 x_p 之间的欧式距离 s_i^{ap} ; 设矢量 $x_i = x_{i1}, x_{i2}, \dots, x_{in}$, 则 s_i^{ap} 计算公式为 $s_i^{ap} = \sqrt{(x_{a1} - x_{p1})^2 + (x_{a2} - x_{p2})^2 + \dots + (x_{an} - x_{pn})^2}$; 随机选取另一个矢量作为 x_n , x_a 和 x_n 分别属于不同的人; 计算 x_a 和 x_n 之间的欧式距离 s_i^{an} ; 则损失函数 L 由下面的公式计算:

$$L = \sum_{i=0}^N \max(s_i^{ap} - s_i^{an} + a, 0); a \text{ 是常量, 取值范围在 } 0.1 \text{ 至 } 0.25 \text{ 之间, } N \text{ 是选取的样例三元组的个数。}$$

参照图 5, 本实施例的所述第一样例和所述第二样例为说话人的话语样例, 所述二分类网络模型为

说话人确认网络模型；输入模块 42，包括：

第一输入单元 421，用于将说话人的话语样例组成的样例三元组训练集输入以 Triplet Loss 为损失函数的说话人确认网络模型中进行模型训练。

本实施例以用于说话人确认为例，人说话时发音器官和发音习惯的差异都以复杂的形势反映在说话人的语音特征中，使得每个人的话语都带有强烈的个人色彩，本实施例通过将话语样例中的语义信息平均化，挖掘出包含在话语样例中的说话人的个性因素，强调不同人之间的语音特征的特征差异。本实施例的批次数据中 M 为 40 个人，R 为 30 个不同语音样例，比如，正常说话状态下选择 10 个不同语音样例，愤怒状态下再选择 10 个不同语音样例，伤心状态下再选择 10 个不同语音样例，将指定人的语音样例与其他人的语音样例组成的样例三元组，本实施例的指定数量为 128，说话人的样例三元组包括说话人的话语样例组成的样例三元组。通过训练说话人确认网络模型训练，形成提取话语样例中的语速、音质、语音波形等语音特征的提取方式，以便通过比较语音特征的相似度，对待测说话人与预注册说话人进行一对一的确认，输出待测说话人与预注册说话人是同一人或待测说话人与预注册说话人不是同一人的确认结果。本实施例通过上述方法在众多说话人的话语样例中，先分成包含指定数量说话人的批次数据，并在批次数据中筛选合适的样例三元组 APN，提高说话人确认网络模型的训练速度以及训练效果，以便扩展说话人确认网络模型的应用领域，包括在金融、证券、社保、公安、军队及其他民用安全认证等行业和部门满足相应的使用需求。

参照图 6，本实施例的第一输入单元 421，包括：

第一输入子单元 4211，用于将待测人的话语样例输入到训练好的所述说话人确认网络模型中。

本实施例的待测人的话语样例包括待测人的自然语音、合成语音等包含音质等一种或多种语音特征的语音样例。举例地，合成语音包括经过去除噪音处理的话语样例、经过去除噪音处理并二次排列拼接的话语样例以及去除噪音后添加了新的背景音等。合成语音中带有强烈的个人色彩的语音特征中至少一种语音特征保持不变，则可实现说话人确认。

第一计算子单元 4212，用于在说话人确认网络模型中计算待测人的话语样例中的第一语音特性与注册人的第二语音特性的第一相似度。

本实施例的待测人的话语样例中的第一语音特性与预注册人的第二语音特性的第一相似度，可通过欧式距离、马氏距离等进行判定，优选通过欧式距离进行判定。本实施例通过 $\max(s_i^{ap} - s_i^{an} + a, 0)$ 筛选候选 N，其中，a 是常量， s_i^{ap} 为 x_a 和 x_p 之间的欧式距离， s_i^{an} 为 x_a 和 x_n 之间的欧式距离； x_a 为话语样例批次数据中的 Anchor 对应的矢量， x_p 为 Positive 对应的矢量； x_n 为 Negative 对应的矢量。通过计算第二矢量距离与所述第一矢量距离的距离差，在距离差大于零的候选 N 集合中，进一步优选距离差最大的候选 N，即优选与指定人的语音特性差别最大的人的语音样例，以提高的样例三元组 APN 的数据代表性，且减少样例三元组 APN 的总数量，以提高说话人确认网络模型训练中的训练速度和训练效果。

第一判断子单元 4213, 用于判断所述第一相似度是否在预设阈值范围内。

举例地, 预设范围为待测人的话语样例中的第一语音特性与预注册人的第二语音特性的欧式距离大于 0.9, 表示第一相似度高或判定第一语音特性的待测人与第二语音特性对应的预注册人为同一个人。

第一判定子单元 4214, 用于若第一相似度在预设阈值范围内, 判定待测人与预注册人为同一个人。

5 参照图 7, 本申请另一实施例的所述第一样例和所述第二样例为人脸影像的样例, 所述二分类网络模型为人脸识别网络模型; 输入模块 42, 包括:

第二输入单元 422, 用于将人脸影像的样例三元组输入以 Triplet Loss 为损失函数的人脸识别网络模型中进行模型训练。

本实施例以用于人脸识别为例, 样例为人脸影像, 用于训练人脸识别模型, 并根据人脸识别模型实
10 现对指定待识别人的人脸影像的一对一判定, 即通过便是结果为是或否, 以便确定待识别人与注册人是否为同一个人。本实施例的批次数据中 M 为 50 个人, R 为 20 个不同角度的人脸影像, 比如, 正面人脸影像旋转到正右侧脸影像之间选择 10 个不同人脸影像, 正面人脸影像旋转到正左侧脸影像之间再选择 10 个不同人脸影像。将指定人的人脸影像与其他人的人脸影像组成的样例三元组。通过对训练人脸识别网络模型训练, 形成提取人脸影像中的人脸特征的提取方式, 以便通过比较人脸特征的相似度, 对待测
15 人与预注册人进行一对一的确认, 输出待测人与预注册人是同一人或待测人与预注册人不是同一人的确认结果。本实施例通过上述方法在 50 个人的人脸影像组成的批次数据中筛选合适的样例三元组 APN, 提高人脸识别网络模型的训练速度以及训练效果, 以便扩展人脸识别网络模型的应用领域, 以便在金融、证券、社保、公安、军队及其他民用安全认证等行业和部门满足使用需求。

参照图 8, 本实施例的第二输入单元 422, 包括:

20 第二输入子单元 4221, 用于将待测人的人脸影像输入到训练好的所述人脸识别网络模型中。

本实施例的待测人的人脸影像包括待测人脸五官的五官特征, 比如五官各自的大小、姿态、位置等。

第二计算子单元 4222, 用于在人脸识别网络模型中计算待测人的人脸影像中的第一人脸特征与预注册人的第二人脸特性的第二相似度。

本实施例通过 $\max(s_i^{ap} - s_i^{an} + a, 0)$ 筛选候选 N, 其中, a 是常量, s_i^{ap} 为 x_a 和 x_p 之间的欧式距离,
25 s_i^{an} 为 x_a 和 x_n 之间的欧式距离; x_a 为人脸影像批次数据中的 Anchor 对应的矢量, x_p 为 Positive 对应的矢量; x_n 为 Negative 对应的矢量。通过计算第二矢量距离与所述第一矢量距离的距离差, 在距离差大于零的候选 N 集合中, 进一步优选距离差最大的候选 N, 即优选与指定人的人脸特性差别最大的人的人脸影像, 以进一步提高的样例三元组 APN 的数据代表性, 且减少样例三元组 APN 的总数量, 以提高人脸识别网络模型训练中的训练速度和训练效果。

30 第二判断子单元 4223, 用于判断所述第二相似度是否在预设阈值范围内。

本实施例的第二相似度可通过欧式距离、马氏距离等进行判定, 优选通过欧式距离进行判定。举例

地, 预设范围为 0.8 或以上, 欧氏距离越大表示相似度高或判定待测人与预注册人为同一个人。

第二判定子单元 4224, 用于若第二相似度在预设阈值范围内, 判定待测人与预注册人为同一个人。

进一步地, 上述第一输入单元 421 或第二输入单元 422, 包括:

控制子单元 4220, 用于控制装配所述说话人确认网络模型的安全认证系统展示通过认证。

- 5 本实施例通过将说话人确认网络模型的确认结果, 转换成与其相关联的安全认证系统的控制指令, 以提高说话人确认系统的实际应用效果。举例地, 说话人确认网络模型的确认待测人与预注册人为同一个人后, 待测人可打开安全认证系统控制的应用平台, 使应用平台处于可使用的状态。本实施例展示的通过认证的方式包括对话框文本显示和/或语音播报和/或警示铃鸣声和/或警示灯的闪光方式等。

- 10 参照图 9, 本申请实施例中还提供一种计算机设备, 该计算机设备可以是服务器, 其内部结构可以如图 9 所示。该计算机设备包括通过系统总线连接的处理器、存储器、网络接口和数据库。其中, 该计算机设计的处理器用于提供计算和控制能力。该计算机设备的存储器包括非易失性存储介质、内存储器。该非易失性存储介质存储有操作系统、计算机可读指令和数据库。该内存储器为非易失性存储介质中的操作系统和计算机可读指令的运行提供环境。该计算机设备的数据库用于存储样例三元组等数据。该计算机设备的网络接口用于与外部的终端通过网络连接通信。该计算机可读指令在执行时, 执行如上述各方法
15 的实施例的流程。本领域技术人员可以理解, 图 9 中示出的结构, 仅仅是与本申请方案相关的部分结构的框图, 并不构成对本申请方案所应用于其上的计算机设备的限定。

- 20 本申请一实施例还提供一种计算机非易失性可读存储介质, 其上存储有计算机可读指令, 该计算机可读指令在执行时, 执行如上述各方法的实施例的流程。以上所述仅为本申请的优选实施例, 并非因此限制本申请的专利范围, 凡是利用本申请说明书及附图内容所作的等效结构或等效流程变换, 或直接或间接运用在其他相关的技术领域, 均同理包括在本申请的专利保护范围内。

权利要求书

1、一种样例三元组的获取方法，其特征在于，包括：

在预设的目标批次数据中筛选属于每一个指定人的所有第一样例，并将所述所有第一样例两两组合以得到每一个指定人对应的多个样例二元组；

5 选择所述多个样例二元组中的第一样例二元组，并计算第一样例二元组中的两个第一样例之间的第一矢量距离，以及分别计算第一样例二元组中指定的目标第一样例与每一个第二样例之间的第二矢量距离；其中，所述第二样例为预设的目标批次数据中除所述第一样例之外的样例数据；

根据每一个第二矢量距离与所述第一矢量距离的差量，筛选最大差量对应的第二样例；

10 将最大差量对应的第二样例与所述第一样例二元组组成所述第一样例二元组对应的多个样例三元组，从而得到所述目标批次数据中的各个指定人对应的每个样例二元组所对应的所有样例三元组。

2、根据权利要求1所述的样例三元组的获取方法，其特征在于，所述在预设的目标批次数据中筛选属于每一个指定人的所有第一样例，并将所述所有第一样例两两组合以得到每一个指定人对应的多个样例二元组的步骤，包括：

将训练数据拆分成包含指定数量样例的多个批次数据；

15 将所述多个批次数据中的目标批次数据输入到预设训练模型中，通过预设训练模型的正向训练获取到所述每个批次数据中所述指定数量样例分别对应的矢量；

筛选所述目标批次数据中属于某一指定人的所有第一样例，并将所述所有第一样例进行两两组合形成多个样例二元组，其中所述第一样例包含与其对应的矢量。

20 3、根据权利要求1所述的样例三元组的获取方法，其特征在于，所述将最大差量对应的所述第二样例与所述第一样例二元组组成所述第一样例二元组对应的多个样例三元组的步骤之后，包括：

将所述训练数据中的所述多个批次数据中所有第一样例二元组分别对应的所有样例三元组汇总成样例三元组训练集；

将所述样例三元组训练集输入以 Triplet Loss 为损失函数的二分类网络模型中进行模型训练。

25 4、根据权利要求3所述的样例三元组的获取方法，其特征在于，所述第一样例和所述第二样例为说话人的话语样例，所述二分类网络模型为说话人确认网络模型；所述将所述样例三元组训练集输入以 Triplet Loss 为损失函数的二分类网络模型中进行模型训练的步骤，包括：

将所述说话人的话语样例组成的样例三元组训练集输入以 Triplet Loss 为损失函数的说话人确认网络模型中进行模型训练。

30 5、根据权利要求4所述的样例三元组的获取方法，其特征在于，所述将所述说话人的话语样例组成的样例三元组训练集输入以 Triplet Loss 为损失函数的说话人确认网络模型中进行模型训练的步骤之后，包括：

将待测人的话语样例输入到训练好的所述说话人确认网络模型中；

在所述说话人确认网络模型中计算所述待测人的话语样例中的第一语音特性与预注册人的第二语音特性的第一相似度；

判断所述第一相似度是否在预设阈值范围内；

5 若是，则判定所述待测人与所述预注册人为同一个人。

6、根据权利要求3所述的样例三元组的获取方法，其特征在于，所述第一样例和所述第二样例为人脸影像的样例，所述二分类网络模型为人脸识别网络模型；所述将所述样例三元组训练集输入以 Triplet Loss 为损失函数的二分类网络模型中进行模型训练的步骤，包括：

10 将所述人脸影像组成的样例三元组训练集输入以 Triplet Loss 为损失函数的人脸识别网络模型中进行模型训练。

7、根据权利要求6所述的样例三元组的获取方法，其特征在于，所述将所述人脸影像组成的样例三元组训练集输入以 Triplet Loss 为损失函数的人脸识别网络模型中进行模型训练的步骤之后，包括：

将待测人的人脸影像输入到训练好的所述人脸识别网络模型中；

15 在所述人脸识别网络模型中计算所述待测人的人脸影像中的第一人脸特性与预注册人的第二人脸特性的第二相似度；

判断所述第二相似度是否在预设阈值范围内；

若是，则判定所述待测人与所述预注册人为同一个人。

8、一种样例三元组的获取装置，其特征在于，包括：

20 第一筛选模块，用于在预设的目标批次数据中筛选属于每一个指定人的所有第一样例，并将所述所有第一样例两两组合以得到每一个指定人对应的多个多个样例二元组；

计算模块，用于选择所述多个样例二元组中的第一样例二元组，并计算第一样例二元组中的两个第一样例之间的第一矢量距离，以及分别计算第一样例二元组中指定的目标第一样例与每一个第二样例之间的第二矢量距离；其中，所述第二样例为预设的目标批次数据中除所述第一样例之外的样例数据；

25 第二筛选模块，用于根据每一个所述第二矢量距离与所述第一矢量距离的差量，筛选最大差量对应的第二样例；

第一组成模块，用于将最大差量对应的第二样例与所述第一样例二元组组成所述第一样例二元组对应的多个样例三元组，从而得到所述目标批次数据中的各个指定人对应的每个样例二元组所对应的所有样例三元组。

9、根据权利要求8所述的样例三元组的获取装置，其特征在于，所述第一筛选模块，包括：

30 拆分单元，用于将训练数据拆分成包含指定数量样例的多个批次数据；

获取单元，用于将所述多个批次数据中的目标批次数据输入到预设训练模型中，通过预设训练模型

的正向训练获取到所述每个批次数据中所述指定数量样例分别对应的矢量；

筛选单元，用于筛选所述目标批次数据中属于某一指定人的所有第一样例，并将所述所有第一样例进行两两组合形成多个样例二元组，其中所述第一样例包含与其对应的矢量。

10、根据权利要求 8 所述的样例三元组的获取装置，其特征在于，包括：

5 汇总模块，用于将所述训练数据中的所述多个批次数据中所有第一样例二元组分别对应的所有样例三元组汇总成样例三元组训练集；

输入模块，用于将所述样例三元组训练集输入以 Triplet Loss 为损失函数的二分类网络模型中进行模型训练。

11、根据权利要求 10 所述的样例三元组的获取装置，其特征在于，所述第一样例和所述第二样例
10 为说话人的话语样例，所述二分类网络模型为说话人确认网络模型；所述输入模块，包括：

第一输入单元，用于将所述说话人的话语样例组成的样例三元组训练集输入以 Triplet Loss 为损失函数的说话人确认网络模型中进行模型训练。

12、根据权利要求 11 所述的样例三元组的获取装置，其特征在于，所述第一输入单元，包括：

第一输入子单元，用于将待测人的话语样例输入到训练好的所述说话人确认网络模型中；

15 第一计算子单元，用于在所述说话人确认网络模型中计算所述待测人的话语样例中的第一语音特性与预注册人的第二语音特性的第一相似度；

第一判断子单元，用于判断所述第一相似度是否在预设阈值范围内；

第一判定子单元，用于，则判定所述待测人与所述预注册人为同一个人。

13、根据权利要求 10 所述的样例三元组的获取装置，其特征在于，所述第一样例和所述第二样例
20 为人脸影像的样例，所述二分类网络模型为人脸识别网络模型；所述输入模块，包括：

第二输入单元，用于将所述人脸影像组成的样例三元组训练集输入以 Triplet Loss 为损失函数的
人脸识别网络模型中进行模型训练。

14、根据权利要求 13 所述的样例三元组的获取装置，其特征在于，所述第二输入单元，包括：

第二输入子单元，用于将待测人的人脸影像输入到训练好的所述人脸识别网络模型中；

25 第二计算子单元，用于在所述人脸识别网络模型中计算所述待测人的人脸影像中的第一人脸特性与预注册人的第二人脸特性的第二相似度；

第二判断子单元，用于判断所述第二相似度是否在预设阈值范围内；

第二判定子单元，用于若第二相似度在预设阈值范围内，则判定所述待测人与所述预注册人为同一个人。

30 15、一种计算机设备，包括存储器和处理器，所述存储器存储有计算机可读指令，其特征在于，所述处理器执行所述计算机可读指令时实现样例三元组的获取方法，方法包括：

在预设的目标批次数据中筛选属于每一个指定人的所有第一样例，并将所述所有第一样例两两组合以得到每一个指定人对应的多个样例二元组；

选择所述多个样例二元组中的第一样例二元组，并计算第一样例二元组中的两个第一样例之间的第一矢量距离，以及分别计算第一样例二元组中指定的目标第一样例与每一个第二样例之间的第二矢量距离；其中，所述第二样例为预设的目标批次数据中除所述第一样例之外的样例数据；

根据每一个第二矢量距离与所述第一矢量距离的差量，筛选最大差量对应的第二样例；

将最大差量对应的第二样例与所述第一样例二元组组成所述第一样例二元组对应的多个样例三元组，从而得到所述目标批次数据中的各个指定人对应的每个样例二元组所对应的所有样例三元组。

16、根据权利要求 15 所述的计算机设备，其特征在于，所述在预设的目标批次数据中筛选属于每一个指定人的所有第一样例，并将所述所有第一样例两两组合以得到每一个指定人对应的多个样例二元组的步骤，包括：

将训练数据拆分成包含指定数量样例的多个批次数据；

将所述多个批次数据中的目标批次数据输入到预设训练模型中，通过预设训练模型的正向训练获取到所述每个批次数据中所述指定数量样例分别对应的矢量；

筛选所述目标批次数据中属于某一指定人的所有第一样例，并将所述所有第一样例进行两两组合形成多个样例二元组，其中所述第一样例包含与其对应的矢量。

17、根据权利要求 15 所述的计算机设备，其特征在于，所述将最大差量对应的所述第二样例与所述第一样例二元组组成所述第一样例二元组对应的多个样例三元组的步骤之后，包括：

将所述训练数据中的所述多个批次数据中所有第一样例二元组分别对应的所有样例三元组汇总成样例三元组训练集；

将所述样例三元组训练集输入以 Triplet Loss 为损失函数的二分类网络模型中进行模型训练。

18、一种计算机非易失性可读存储介质，其上存储有计算机可读指令，其特征在于，所述计算机可读指令被处理器执行时实现样例三元组的获取方法，方法包括：

在预设的目标批次数据中筛选属于每一个指定人的所有第一样例，并将所述所有第一样例两两组合以得到每一个指定人对应的多个样例二元组；

选择所述多个样例二元组中的第一样例二元组，并计算第一样例二元组中的两个第一样例之间的第一矢量距离，以及分别计算第一样例二元组中指定的目标第一样例与每一个第二样例之间的第二矢量距离；其中，所述第二样例为预设的目标批次数据中除所述第一样例之外的样例数据；

根据每一个第二矢量距离与所述第一矢量距离的差量，筛选最大差量对应的第二样例；

将最大差量对应的第二样例与所述第一样例二元组组成所述第一样例二元组对应的多个样例三元组，从而得到所述目标批次数据中的各个指定人对应的每个样例二元组所对应的所有样例三元组。

19、根据权利要求 18 所述的计算机非易失性可读存储介质，其特征在于，所述在预设的目标批次数据中筛选属于每一个指定人的所有第一样例，并将所述所有第一样例两两组合以得到每一个指定人对应的多个样例二元组的步骤，包括：

将训练数据拆分成包含指定数量样例的多个批次数据；

5 将所述多个批次数据中的目标批次数据输入到预设训练模型中，通过预设训练模型的正向训练获取到所述每个批次数据中所述指定数量样例分别对应的矢量；

筛选所述目标批次数据中属于某一指定人的所有第一样例，并将所述所有第一样例进行两两组合形成多个样例二元组，其中所述第一样例包含与其对应的矢量。

20、根据权利要求 18 所述的计算机非易失性可读存储介质，其特征在于，所述将最大差量对应的
10 所述第二样例与所述第一样例二元组组成所述第一样例二元组对应的多个样例三元组的步骤之后，包括：

将所述训练数据中的所述多个批次数据中所有第一样例二元组分别对应的所有样例三元组汇总成样例三元组训练集；

将所述样例三元组训练集输入以 Triplet Loss 为损失函数的二分类网络模型中进行模型训练。

15

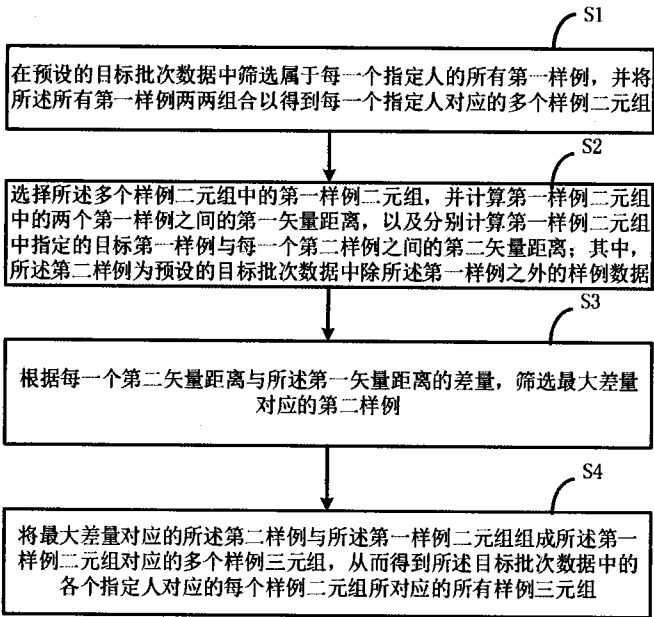


图 1

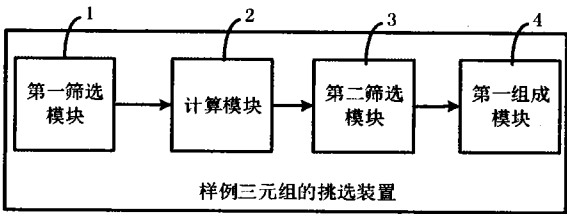


图 2

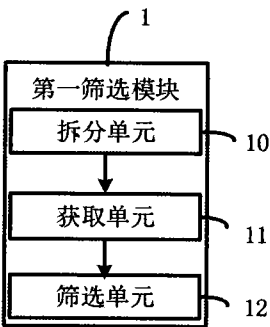


图 3

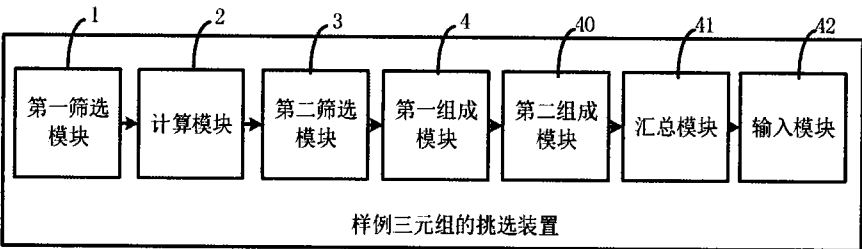


图 4

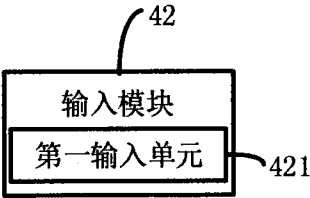


图 5

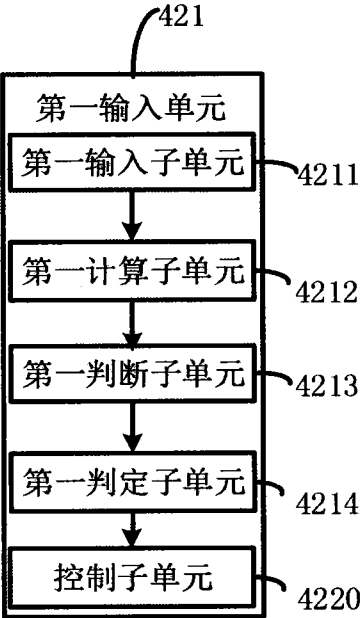


图 6

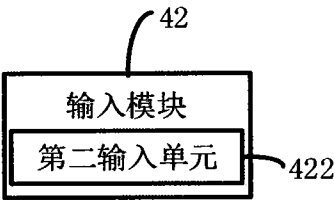


图 7

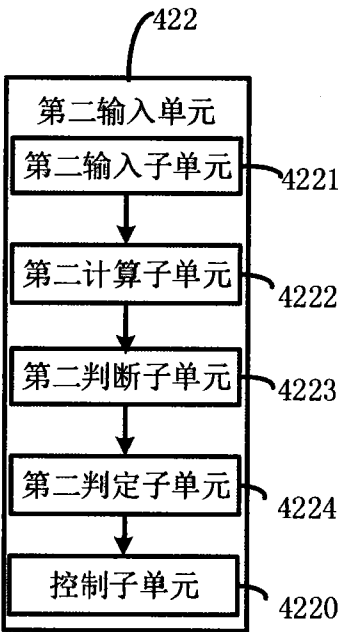


图 8

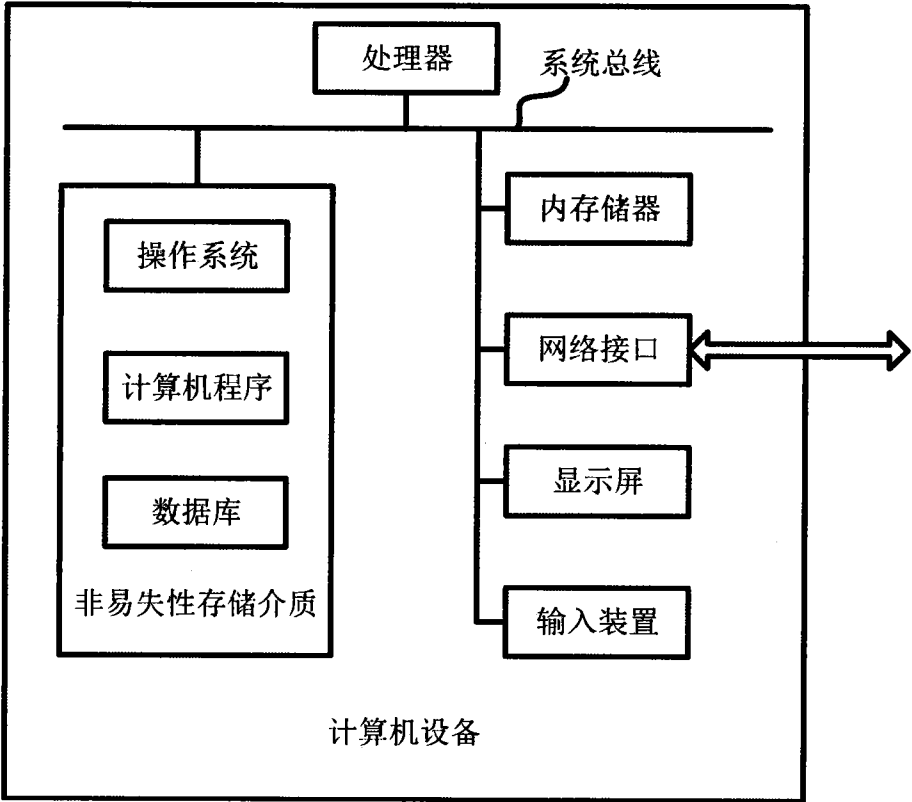


图 9

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2018/095462

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/00(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06K, G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS, CNKI, DWPI, SIPOABS: 样例, 实例, 元组, 组合, 选择, 选取, 筛选, 矢量, 向量, 距离, 差, example, sample, instance, tuple, combination, group, select, choose, elect, filter, filtrate, vector, distance, difference

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 106484675 A (BEIJING INSTITUTE OF TECHNOLOGY) 08 March 2017 (2017-03-08) entire document	1-20
A	CN 105740327 A (TIANJIN ZHONGKE INSTITUTE OF INTELLIGENT IDENTIFICATION INDUSTRY TECHNOLOGY CO., LTD.) 06 July 2016 (2016-07-06) entire document	1-20
A	US 2013212090 A1 (STROZ FRIEDBERG, LLC) 15 August 2013 (2013-08-15) entire document	1-20
A	US 2006074604 A1 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 06 April 2006 (2006-04-06) entire document	1-20

☐ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

15 January 2019

Date of mailing of the international search report

30 January 2019

Name and mailing address of the ISA/CN

**National Intellectual Property Administration, PRC (ISA/
CN)**
**No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing
100088
China**

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2018/095462

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	106484675	A	08 March 2017	None			
CN	105740327	A	06 July 2016	None			
US	2013212090	A1	15 August 2013	US	9208219	B2	08 December 2015
US	2006074604	A1	06 April 2006	US	7392161	B2	24 June 2008
				US	2007106485	A9	10 May 2007

国际检索报告

国际申请号

PCT/CN2018/095462

A. 主题的分类

G06K 9/00(2006.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G06K, G10L

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

CNABS, CNKI, DWPI, SIPOABS: 样例, 实例, 元组, 组合, 选择, 选取, 筛选, 矢量, 向量, 距离, 差, example, sample, instance, tuple, combination, group, select, choose, elect, filter, filtrate, vector, distance, difference

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
A	CN 106484675 A (北京理工大学) 2017年 3月 8日 (2017 - 03 - 08) 全文	1-20
A	CN 105740327 A (天津中科智能识别产业技术研究院有限公司) 2016年 7月 6日 (2016 - 07 - 06) 全文	1-20
A	US 2013212090 A1 (STROZ FRIEDBERG LLC) 2013年 8月 15日 (2013 - 08 - 15) 全文	1-20
A	US 2006074604 A1 (INTERNATIONAL BUSINESS MACHINES IBM CORPORATION) 2006年 4月 6日 (2006 - 04 - 06) 全文	1-20

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2019年 1月 15日

国际检索报告邮寄日期

2019年 1月 30日

ISA/CN的名称和邮寄地址

中国国家知识产权局(ISA/CN)
中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

受权官员

邓隽

电话号码 86-(10)-62411644

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2018/095462

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	106484675	A	2017年 3月 8日	无			
CN	105740327	A	2016年 7月 6日	无			
US	2013212090	A1	2013年 8月 15日	US	9208219	B2	2015年 12月 8日
US	2006074604	A1	2006年 4月 6日	US	7392161	B2	2008年 6月 24日
				US	2007106485	A9	2007年 5月 10日

表 PCT/ISA/210 (同族专利附件) (2015年1月)