



(12)发明专利申请

(10)申请公布号 CN 108154066 A  
(43)申请公布日 2018.06.12

(21)申请号 201611096314.1

(22)申请日 2016.12.02

(71)申请人 中国科学院沈阳自动化研究所  
地址 110016 辽宁省沈阳市东陵区南塔街  
114号

(72)发明人 梁炜 李杨 郑萌 谈金东  
彭士伟

(74)专利代理机构 沈阳科苑专利商标代理有限公司 21002

代理人 许宗富

(51)Int.Cl.  
G06K 9/00(2006.01)  
G06N 3/04(2006.01)

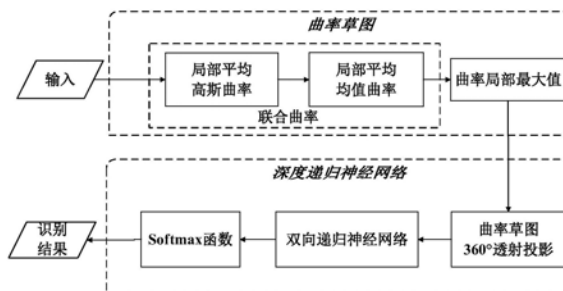
权利要求书2页 说明书5页 附图1页

(54)发明名称

一种基于曲率特征递归神经网络的三维目标识别方法

(57)摘要

本发明涉及图像识别技术,为了有效地刻画三维目标在不同视角下的特征,针对三维目标识别过程中存在的图像噪声问题,提出了一种基于曲率特征递归神经网络的三维目标识别方法。首先,本发明通过计算目标三维模型的局部平均高斯曲率和平均均值曲率得出目标三维模型的联合曲率,并通过提取联合曲率局部极大值构成三维模型的曲率草图,利用透射投影变换生成360°二维图像序列作为训练递归神经网络的输入;其次,利用双向递归神经网络(BRNN)作为三维模型多视角序列特征学习方法,在softmax层利用softmax函数求得正确概率最大的识别类别。本发明能够自动提取三维目标与二维图像的共同特征,能够在图像噪声条件下保持较好的鲁棒性和较高的目标识别率。



1. 一种基于曲率特征递归神经网络的三维目标识别方法,其特征在于,包括以下步骤:

步骤1:计算目标三维模型的联合曲率 $\bar{U}_K$ ,提取联合曲率 $\bar{U}_K$ 的局部极大值构成三维模型的曲率草图 $R_{Sketch}$ ;再对三维模型的曲率草图 $R_{Sketch}$ 利用透射投影变换生成 $360^\circ$ 二维图像 $P_m$ ,其中 $m=1,2,\dots,360$ ;

步骤2:将 $360^\circ$ 二维图像输入BRNN,利用多角度特征进行学习计算其在多视角下的序列属性;在softmax层利用softmax函数求得序列属性的正确概率最大时的识别类别;所述BRNN为双向递归神经网络。

2. 按照权利要求1所述一种基于曲率特征递归神经网络的三维目标识别方法,其特征在于,所述计算目标三维模型的联合曲率 $\bar{U}_K$ 包括以下步骤:

设 $\vec{n} = (n_x, n_y, n_z)'$ 是目标三维模型R上给定一点 $(x, y, z)$ 的法向量;令 $p = \frac{n_x}{n_z}, q = \frac{n_y}{n_z}$ ,则 $p_x, p_y, q_x, q_y$ 定义为 $p_x = \frac{\partial p}{\partial x}, p_y = \frac{\partial p}{\partial y}, q_x = \frac{\partial q}{\partial x}, q_y = \frac{\partial q}{\partial y}$ ;

计算三维模型R上每一点的法向量周围 $3 \times 3$ 邻域内的平均高斯曲率 $\bar{G}_K$ 和平均均值曲率 $\bar{M}_K$ :

$$\bar{G}_K = |\bar{C}|,$$

$$\bar{M}_K = \frac{1}{2} \text{trace}(\bar{C}),$$

其中, $\bar{C} = (\bar{p}^2 + \bar{q}^2 + 1)^{-3/2} \begin{pmatrix} \bar{q}^2 + 1 & -\bar{p}\bar{q} \\ -\bar{p}\bar{q} & \bar{p}^2 + 1 \end{pmatrix} \begin{pmatrix} \bar{p}_x & \bar{p}_y \\ \bar{q}_x & \bar{q}_y \end{pmatrix}$ 为平均曲率矩阵,  $\text{trace}(\cdot)$

是矩阵的迹,  $\bar{p}, \bar{q}, \bar{p}_x, \bar{p}_y, \bar{q}_x, \bar{q}_y$ 分别为 $p, q, p_x, p_y, q_x, q_y$ 在 $3 \times 3$ 邻域内的平均值;

定义目标三维模型R的联合曲率 $\bar{U}_K$ 为:

$$\bar{U}_K = \left| \frac{\bar{G}_K - \bar{M}_K}{\bar{G}_K + \bar{M}_K} \right|.$$

3. 按照权利要求1所述一种基于曲率特征递归神经网络的三维目标识别方法,其特征在于,所述将 $360^\circ$ 二维图像输入BRNN,利用多角度特征进行学习计算出其在多视角下的序列属性,包括以下步骤:

获取 $360^\circ$ 二维图像的一维特征序列 $T_s, s=1,2,\dots,360$ ,则特征序列 $T_s$ 在BRNN第i层的输出分为正向输出 $\vec{O}_i^s$ 和反向输出 $\vec{O}_i^s$ ,并且分别与本层BRNN上一序列的正向输出 $\vec{O}_{i-1}^{s-1}$ 、本层BRNN下一序列的反向输出 $\vec{O}_i^{s+1}$ 以及上一层BRNN的正向输出 $\vec{O}_{i-1}^s$ 和反向输出 $\vec{O}_{i-1}^s$ 有如下关系:

$$\vec{O}_i^s = \tanh \left( W_{\vec{O}_{i-1}, \vec{O}_i} \vec{O}_{i-1}^s + W_{\vec{O}_{i-1}, \vec{O}_i} \vec{O}_{i-1}^s + W_{\vec{O}_i, \vec{O}_i} \vec{O}_i^{s-1} + b_{\vec{O}_i} \right),$$

$$\vec{O}_i^s = \tanh \left( W_{\vec{O}_{i-1}, \vec{O}_i} \vec{O}_{i-1}^s + W_{\vec{O}_{i-1}, \vec{O}_i} \vec{O}_{i-1}^s + W_{\vec{O}_i, \vec{O}_i} \vec{O}_i^{s+1} + b_{\vec{O}_i} \right),$$

其中,  $W = \begin{pmatrix} W_{\vec{o}_{i-1}, \vec{o}_i} & W_{\vec{o}_{i-1}, \vec{o}_i} & W_{\vec{o}_i, \vec{o}_i} \\ W_{\vec{o}_{i-1}, \vec{o}_i} & W_{\vec{o}_{i-1}, \vec{o}_i} & W_{\vec{o}_i, \vec{o}_i} \end{pmatrix}$  为各输出间的权值矩阵,  $b$  为偏置,  $\tanh$  为神经

元激活函数;

则特征序列  $T_s$  在 BRNN 的总输出  $O^s$ , 即为全连接层  $f_c$  的输入  $I_{f_c}$  为:

$$O^s = I_{f_c} = W_{\vec{o}, f_c} \vec{O}_i + W_{\overleftarrow{o}, f_c} \overleftarrow{O}_i;$$

其中,  $W_{\vec{o}, f_c}$ 、 $W_{\overleftarrow{o}, f_c}$  分别为正向输出和反向输出在全连接层的连接权值;

因此, 特征序列  $T_s$  在全连接层  $f_c$  的累加输出为  $A = \sum_0^{s-1} O^s$ , 即为序列属性。

4. 按照权利要求 1 所述一种基于曲率特征递归神经网络的三维目标识别方法, 其特征在于, 所述在 softmax 层利用 softmax 函数求得序列属性的正确概率最大时的识别类别, 包括以下步骤:

在 softmax 层利用 softmax 函数计算出识别结果为第  $k$  类的正确概率  $p(C_k)$

$$p(C_k) = \frac{e^{A_k}}{\sum_{k=0}^{C-1} e^{A_k}},$$

其中,  $C$  为识别类别总数,  $A_k$  为第  $k$  类三维目标的序列属性在全连接层  $f_c$  的累加输出结果;

然后利用最大似然估计方法求得损失函数最小值时, 即正确概率  $p(C_k)$  最大时的识别类别  $k$ :

$$L(s) = -\ln \sum_{k=0}^{C-1} \delta(k-r) p(C_k | s),$$

其中,  $\delta(\cdot)$  是克罗内克函数  $\delta(k-r) = \begin{cases} 0, & \text{当 } k \neq r \text{ 时} \\ 1, & \text{当 } k = r \text{ 时} \end{cases}$ ,  $r$  表示特征序列  $T_s$  的正确识别

类别。

## 一种基于曲率特征递归神经网络的三维目标识别方法

### 技术领域

[0001] 本发明涉及图像识别技术领域,具体地说是一种基于曲率特征递归神经网络的三维目标识别方法。

### 背景技术

[0002] 三维目标识别是指从任意给定的二维图像场景中自动检测、定位、识别出指定目标模式的过程,是计算机视觉研究的关键问题之一。随着计算机视觉技术的不断发展,三维目标识别越来越广泛地应用于工业检测、增强现实和医学影像等领域。但是,由于受到光照变化、图像噪声和目标遮挡等因素的影响,难以提取三维目标及其在不同视角下二维图像的共同特征,成为三维目标识别亟待解决的问题。

[0003] 三维目标识别的关键是找到三维目标模型的二维表达,提取三维目标和二维图像的共同特征。现有的三维目标识别方法主要包括基于人工标记点的方法、基于几何特征的方法和基于深度学习的方法等。基于人工标记点的方法需要人工初始化二维图像中的特征点,由于需要人工交互,所以此类方法不具有可重复性;基于几何特征的方法通过提取目标的中线骨架、轮廓形状等信息实现目标识别,但是此类方法在图像存在噪声的情况下识别效果较差;基于深度学习的方法利用深度神经网络将低水平的图像特征融合成带有语义信息的高水平特征,能够很好地解决三维目标识别过程中二维图像的图像噪声问题,但是通常使用的深度卷积神经网络无法表达序列属性,不能有效地刻画三维目标在不同视角下的特征。因此,亟需提出一种在不同视角图像中对图像噪声问题鲁棒的自动化三维目标识别方法。

### 发明内容

[0004] 本发明目的是能够更有效地刻画三维目标在不同视角下的特征,降低特征提取过程对图像噪声的敏感程度,提高三维目标识别准确率,本发明提出一种基于曲率特征递归神经网络的三维目标识别方法。

[0005] 本发明为实现上述目的所采用的技术方案是:一种基于曲率特征递归神经网络的三维目标识别方法,包括以下步骤:

[0006] 步骤1:计算目标三维模型的联合曲率 $\bar{U}_k$ ,提取联合曲率 $\bar{U}_k$ 的局部极大值构成三维模型的曲率草图 $R_{Sketch}$ ;再对三维模型的曲率草图 $R_{Sketch}$ 利用透射投影变换生成 $360^\circ$ 二维图像 $P_m$ ,其中 $m=1,2,\dots,360$ ;

[0007] 步骤2:将 $360^\circ$ 二维图像输入BRNN,利用多角度特征进行学习计算其在多视角下的序列属性;在softmax层利用softmax函数求得序列属性的正确概率最大时的识别类别;所述BRNN为双向递归神经网络。

[0008] 所述计算目标三维模型的联合曲率 $\bar{U}_k$ 包括以下步骤:

[0009] 设 $\vec{n} = (n_x, n_y, n_z)'$ 是目标三维模型 $R$ 上给定一点 $(x, y, z)$ 的法向量;令

$p = \frac{n_x}{n_z}, q = \frac{n_y}{n_z}$ , 则  $p_x, p_y, q_x, q_y$  定义为  $p_x = \frac{\partial p}{\partial x}, p_y = \frac{\partial p}{\partial y}, q_x = \frac{\partial q}{\partial x}, q_y = \frac{\partial q}{\partial y}$ ;

[0010] 计算三维模型R上每一点的法向量周围  $3 \times 3$  邻域内的平均高斯曲率  $\bar{G}_K$  和平均均值曲率  $\bar{M}_K$ :

$$[0011] \quad \bar{G}_K = |\bar{C}|,$$

$$[0012] \quad \bar{M}_K = \frac{1}{2} \text{trace}(\bar{C}),$$

[0013] 其中,  $\bar{C} = (\bar{p}^2 + \bar{q}^2 + 1)^{-3/2} \begin{pmatrix} \bar{q}^2 + 1 & -\bar{p}\bar{q} \\ -\bar{p}\bar{q} & \bar{p}^2 + 1 \end{pmatrix} \begin{pmatrix} \bar{p}_x & \bar{p}_y \\ \bar{q}_x & \bar{q}_y \end{pmatrix}$  为平均曲率矩阵,

$\text{trace}(\cdot)$  是矩阵的迹,  $\bar{p}, \bar{q}, \bar{p}_x, \bar{p}_y, \bar{q}_x, \bar{q}_y$  分别为  $p, q, p_x, p_y, q_x, q_y$  在  $3 \times 3$  邻域内的平均值;

[0014] 定义目标三维模型R的联合曲率  $\bar{U}_K$  为:

$$[0015] \quad \bar{U}_K = \left| \frac{\bar{G}_K - \bar{M}_K}{\bar{G}_K + \bar{M}_K} \right|.$$

[0016] 所述将  $360^\circ$  二维图像输入BRNN, 利用多角度特征进行学习计算出其在多视角下的序列属性, 包括以下步骤:

[0017] 获取  $360^\circ$  二维图像的一维特征序列  $T_s, s=1, 2, \dots, 360$ , 则特征序列  $T_s$  在BRNN第  $i$  层的输出分为正向输出  $\vec{O}_i^s$  和反向输出  $\bar{O}_i^s$ , 并且分别与本层BRNN上一序列的正向输出  $\vec{O}_i^{s-1}$ 、本层BRNN下一序列的反向输出  $\bar{O}_i^{s+1}$  以及上一层BRNN的正向输出  $\vec{O}_{i-1}^s$  和反向输出  $\bar{O}_{i-1}^s$  有如下关系:

$$[0018] \quad \vec{O}_i^s = \tanh \left( W_{\vec{O}_{i-1}, \vec{O}_i} \vec{O}_{i-1}^s + W_{\vec{O}_{i-1}, \bar{O}_i} \bar{O}_{i-1}^s + W_{\vec{O}_i, \vec{O}_i} \vec{O}_i^{s-1} + b_{\vec{O}_i} \right),$$

$$[0019] \quad \bar{O}_i^s = \tanh \left( W_{\bar{O}_{i-1}, \vec{O}_i} \vec{O}_{i-1}^s + W_{\bar{O}_{i-1}, \bar{O}_i} \bar{O}_{i-1}^s + W_{\bar{O}_i, \bar{O}_i} \bar{O}_i^{s+1} + b_{\bar{O}_i} \right),$$

[0020] 其中,  $W = \begin{pmatrix} W_{\vec{O}_{i-1}, \vec{O}_i} & W_{\vec{O}_{i-1}, \bar{O}_i} & W_{\vec{O}_i, \vec{O}_i} \\ W_{\bar{O}_{i-1}, \vec{O}_i} & W_{\bar{O}_{i-1}, \bar{O}_i} & W_{\bar{O}_i, \bar{O}_i} \end{pmatrix}$  为各输出间的权值矩阵,  $b$  为偏置,  $\tanh$  为

神经元激活函数;

[0021] 则特征序列  $T_s$  在BRNN的总输出  $O^s$ , 即为全连接层  $fc$  的输入  $I_{fc}$  为:

$$[0022] \quad O^s = I_{fc} = W_{\vec{O}, fc} \vec{O}_i + W_{\bar{O}, fc} \bar{O}_i;$$

[0023] 其中,  $W_{\vec{O}, fc}$ 、 $W_{\bar{O}, fc}$  分别为正向输出和反向输出在全连接层的连接权值;

[0024] 因此, 特征序列  $T_s$  在全连接层  $fc$  的累加输出为  $A = \sum_0^{s-1} O^s$ , 即为序列属性。

[0025] 所述在softmax层利用softmax函数求得序列属性的正确概率最大时的识别类别,

包括以下步骤:

[0026] 在softmax层利用softmax函数计算出识别结果为第k类的正确概率 $p(C_k)$

$$[0027] \quad p(C_k) = \frac{e^{A_k}}{\sum_{k=0}^{C-1} e^{A_k}},$$

[0028] 其中,C为识别类别总数, $A_k$ 为第k类三维目标的序列属性在全连接层fc的累加输出结果;

[0029] 然后利用最大似然估计方法求得损失函数最小值时,即正确概率 $p(C_k)$ 最大时的识别类别k:

$$[0030] \quad L(s) = -\ln \sum_{k=0}^{C-1} \delta(k-r) p(C_k | s),$$

[0031] 其中, $\delta(\cdot)$ 是克罗内克函数 $\delta(k-r) = \begin{cases} 0, & \text{当 } k \neq r \text{ 时} \\ 1, & \text{当 } k = r \text{ 时} \end{cases}$ ,r表示特征序列 $T_s$ 的正

确识别类别。

[0032] 本发明具有以下有益效果及优点:

[0033] 1.本发明设计的联合曲率草图特征提取方法,能够自动提取三维模型与二维图像的共同特征,并且联合曲率所使用的局部平均高斯曲率和局部平均均值曲率可以有效的解决图像噪声问题。

[0034] 2.本发明设计多角度特征学习双向递归神经网络,能够同时考虑三维模型在多角度下的特征序列,能够在任意角度的二维图像中准确识别三维目标。

## 附图说明

[0035] 图1为本发明方法流程图;

[0036] 图2为本发明方法中的多角度特征学习双向递归神经网络框架图。

## 具体实施方式

[0037] 下面结合附图及实施例对本发明做进一步的详细说明。

[0038] 本发明主要分为两部分,如图1所示为本发明方法流程图,具体实现过程如下所述。

[0039] 步骤1:计算目标三维模型的联合曲率,并通过提取联合曲率局部极大值构成三维模型的曲率草图,利用透射投影变换生成 $360^\circ$ 二维图像作为训练递归神经网络的输入;

[0040] 步骤1.1:设 $\vec{n} = (n_x, n_y, n_z)'$ 是三维模型上给定一点 $(x, y, z)$ 的法向量。令

$p = \frac{n_x}{n_z}, q = \frac{n_y}{n_z}$ ,则 $p_x, p_y, q_x, q_y$ 定义为 $p_x = \frac{\partial p}{\partial x}, p_y = \frac{\partial p}{\partial y}, q_x = \frac{\partial q}{\partial x}, q_y = \frac{\partial q}{\partial y}$ ,则三维模型

的高斯曲率 $G_K$ 为

[0041]  $G_K = |C|$ ,

[0042] 其中曲率矩阵  $C = (p^2 + q^2 + 1)^{-3/2} \begin{pmatrix} q^2 + 1 & -pq \\ -pq & p^2 + 1 \end{pmatrix} \begin{pmatrix} p_x & p_y \\ q_x & q_y \end{pmatrix}$ , 三维模型的

均值曲率  $M_K$  为  $M_K = \frac{1}{2} \text{trace}(C)$ ,  $\text{trace}(\cdot)$  是矩阵的迹。为了消除噪声影响, 本发明计算三维模型上每一点的法向量周围其  $3 \times 3$  邻域内的平均高斯曲率  $\bar{G}_K$  和平均均值曲率  $\bar{M}_K$ :

$$[0043] \quad \bar{G}_K = |\bar{C}|,$$

$$[0044] \quad \bar{M}_K = \frac{1}{2} \text{trace}(\bar{C}),$$

[0045] 其中  $\bar{C} = (\bar{p}^2 + \bar{q}^2 + 1)^{-3/2} \begin{pmatrix} \bar{q}^2 + 1 & -\bar{p}\bar{q} \\ -\bar{p}\bar{q} & \bar{p}^2 + 1 \end{pmatrix} \begin{pmatrix} \bar{p}_x & \bar{p}_y \\ \bar{q}_x & \bar{q}_y \end{pmatrix}$  为平均曲率矩阵,

$\bar{p}, \bar{q}, \bar{p}_x, \bar{p}_y, \bar{q}_x, \bar{q}_y$  分别为  $p, q, p_x, p_y, q_x, q_y$  在  $3 \times 3$  邻域内的平均值。由此, 我们定义三维模型的联合曲率  $\bar{U}_K$  为

$$[0046] \quad \bar{U}_K = \left| \frac{\bar{G}_K - \bar{M}_K}{\bar{G}_K + \bar{M}_K} \right|。$$

[0047] 步骤1.2: 提取联合曲率  $\bar{U}_K$  的局部最大值点构成三维模型  $R$  的曲率草图  $R_{\text{Sketch}}$ 。通过透视投影变换, 生成三维曲率草图  $R_{\text{Sketch}}$  的  $360^\circ$  二维投影图像  $P_m, m=1, 2, \dots, 360$ , 作为 BRNN 的输入。

[0048] 步骤2: 本发明采用一种深度递归神经网络 (DRNN) 作为曲率特征识别方法, DRNN 框架如图2所示。利用多角度特征学习 BRNN 刻画三维模型在多视角下的序列属性, 在 softmax 层利用 softmax 函数求得正确概率最大的识别类别。

[0049] 步骤2.1: 为了刻画三维模型在不同视角下特征的序列性, 定义三维模型在多视角下的一维特征序列为  $T_s, s=1, 2, \dots, 360$ , 则特征序列  $T_s$  在 BRNN 第  $i$  层的输出分为正向输出  $\vec{O}_i^s$  和反向输出  $\vec{O}_i^s$ , 分别与本层 BRNN 上一序列的正向输出  $\vec{O}_{i-1}^{s-1}$ 、本层 BRNN 下一序列的反向输出  $\vec{O}_i^{s+1}$  以及上一层 BRNN 的正向输出  $\vec{O}_{i-1}^s$  和反向输出  $\vec{O}_{i-1}^s$  有如下关系:

$$[0050] \quad \vec{O}_i^s = \tanh \left( W_{\vec{O}_{i-1}, \vec{O}_i} \vec{O}_{i-1}^s + W_{\vec{O}_{i-1}, \vec{O}_i} \vec{O}_{i-1}^s + W_{\vec{O}_i, \vec{O}_i} \vec{O}_i^{s-1} + b_{\vec{O}_i} \right),$$

$$[0051] \quad \vec{O}_i^s = \tanh \left( W_{\vec{O}_{i-1}, \vec{O}_i} \vec{O}_{i-1}^s + W_{\vec{O}_{i-1}, \vec{O}_i} \vec{O}_{i-1}^s + W_{\vec{O}_i, \vec{O}_i} \vec{O}_i^{s+1} + b_{\vec{O}_i} \right),$$

[0052] 其中  $W = \begin{pmatrix} W_{\vec{O}_{i-1}, \vec{O}_i} & W_{\vec{O}_{i-1}, \vec{O}_i} & W_{\vec{O}_i, \vec{O}_i} \\ W_{\vec{O}_{i-1}, \vec{O}_i} & W_{\vec{O}_{i-1}, \vec{O}_i} & W_{\vec{O}_i, \vec{O}_i} \end{pmatrix}$  为各输出间的权值矩阵,  $b$  为偏执,  $\tanh$  为神经元激活函数; 则特征序列  $T_s$  在 BRNN 的总输出  $O^s$ , 即为全连接层  $fc$  的输入  $I_{fc}$  为

$$[0053] \quad O^s = I_{fc} = W_{\vec{O}, fc} \vec{O}_i + W_{\vec{O}, fc} \vec{O}_i,$$

[0054] 其中,  $W_{\bar{o},fc}$ 、 $W_{\bar{o},fc}$ 分别为正向输出和反向输出在全连接层的连接权值。

[0055] 步骤2.2:特征序列 $T_s$ 在全连接层 $fc$ 的累加输出为 $A = \sum_0^{s-1} O^s$ ,即为序列属性。在softmax层利用softmax函数计算识别结果为第 $k$ 类的正确概率 $p(C_k)$

$$[0056] \quad p(C_k) = \frac{e^{A_k}}{\sum_{k=0}^{C-1} e^{A_k}},$$

[0057] 其中 $C$ 为识别类别总数, $A_k$ 为第 $k$ 类三维目标的序列属性在全连接层 $fc$ 的累加输出结果。然后利用最大似然估计方法求得损失函数最小值时,即正确概率 $p(C_k)$ 最大时的识别类别 $k$ :

$$[0058] \quad L(s) = -\ln \sum_{k=0}^{C-1} \delta(k-r) p(C_k | s),$$

[0059] 其中 $\delta(\cdot)$ 是克罗内克函数 $\delta(k-r) = \begin{cases} 0, & \text{if } k \neq r \\ 1, & \text{if } k = r \end{cases}$ , $r$ 表示特征序列 $T_s$ 的正确识别类别。



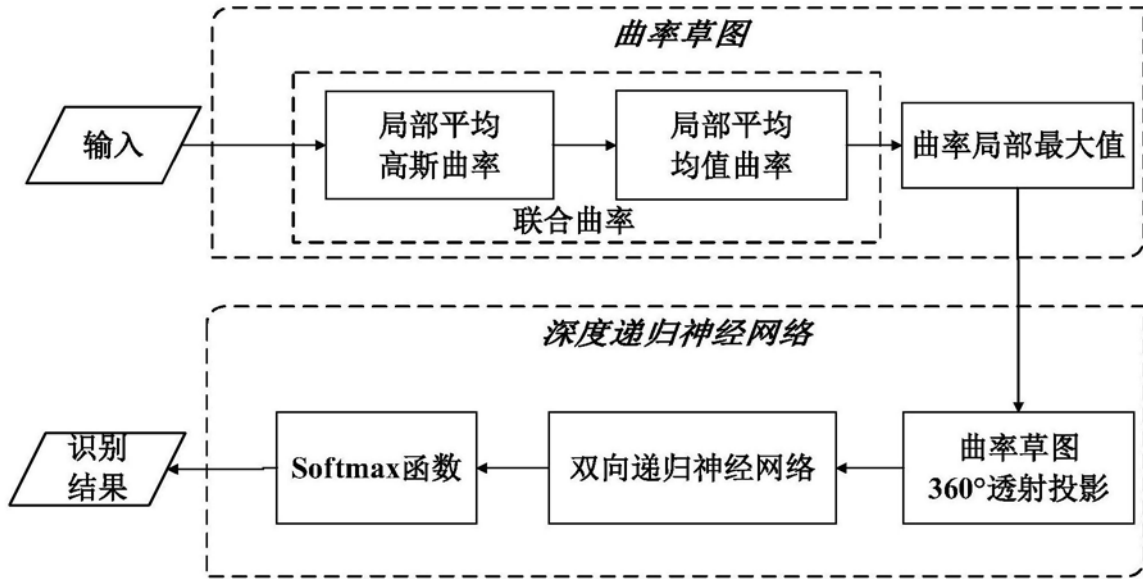


图1

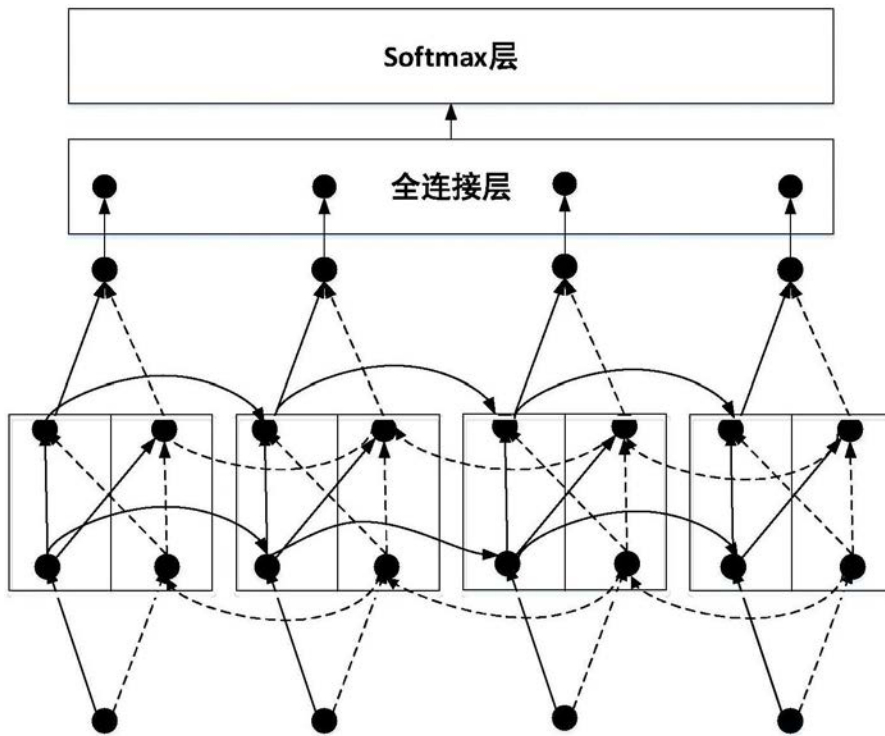


图2