(19)国家知识产权局

(12)发明专利

(10)授权公告号 CN 112740266 B
(45)授权公告日 2024.09.13

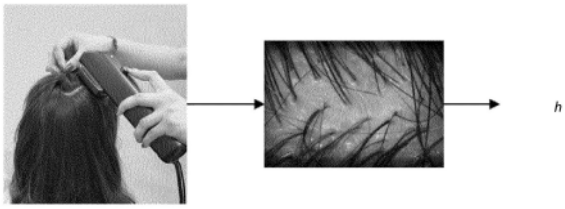
(21)申请号 201980061174.X
(22)申请日 2019.07.15
(65)同一申请的已公布的文献号
申请公布号 CN 112740266 A
(43)申请公布日 2021.04.30
(30)优先权数据
18195776.2 2018.09.20 EP
(85)PCT国际申请进入国家阶段日
2021.03.18
(86)PCT国际申请的申请数据
PCT/EP2019/069047 2019.07.15
(87)PCT国际申请的公布数据
WO2020/057803 EN 2020.03.26
(73)专利权人 莱雅公司
地址 法国巴黎
(72)发明人 M•佩罗特 E•马勒布
T•瓦塞尔曼 J•沙尔贝
P•博卡里斯
(74)专利代理机构 中国专利代理(香港)有限公司 72001
专利代理师 郭慧 林毅斌
(51)Int.Cl.
G06T 7/00(2017.01)
(56)对比文件
US 2007058858 A1,2007.03.15
US 2018014777 A1,2018.01.18
Nazia Hameed等.A comprehensive survey on image-based computer aided diagnosis systems for skin cancer.2016 10th International Conference on Software, Knowledge, Information Management & Applications.2016,第205-241页.
审查员 佟晓明
权利要求书1页 说明书11页 附图2页

(54)发明名称

用于确定角蛋白表面的特性的方法和系统以及用于处理所述角蛋白表面的方法和系统

(57)摘要

本申请涉及一种用于确定使用者的角蛋白表面的至少一个物理和/或化学特性的方法和系统,所述方法包括以下步骤:-接收与所述角蛋白表面的至少一个图像相对应的数据,-通过对所述图像应用至少一个机器学习模型来处理所述图像,-返回与所述角蛋白表面的要确定的特性的评分等级相对应的至少一个数值。



CN 112740266 B

1.一种用于确定使用者的头发色调的计算机执行方法,该方法包括以下步骤:

-接收与以RGB或CMYK或另一颜色空间编码的角蛋白表面的至少一个图像对应的数据,所述图像图示了使用者的包括头发的头皮区域,所述图像是以1:1或更高的放大率拍摄的放大图像,

-应用被配置为识别至少一个头发分段的预处理分割步骤,

-通过应用机器学习模型将至少一个识别的头发分段的颜色值转换为CIELAB颜色空间来处理该图像,然后连续应用至少一个经过相应训练的第二个机器学习模型以由头发像素的中位L*值评估所述头发色调-返回与所述角蛋白表面的要确定的特性的评分等级相对应的至少一个数值。

2.根据权利要求1所述的方法,其特征在于,所述分割步骤通过对比分析来执行。

3.根据权利要求1所述的方法,其特征在于,通过将机器学习模型应用于所述图像来执行所述分割步骤。

4.根据权利要求1至3中任一项所述的方法,其特征在于,所述图像是RGB图像。

5.根据权利要求1至4中任一项所述的方法,其特征在于,重复所述方法以确定所述角蛋白表面的第二特性,优选地使用相同的图像数据重复所述方法。

6.根据权利要求1至5中任一项所述的方法,其特征在于,所述方法包括基于至少一些实值特性来推荐化妆品产品的后续步骤。

7.根据权利要求1至5中任一项所述的方法,其特征在于,所述方法包括基于至少一些实值特性来确定推荐的化妆品产品组合物的后续步骤。

8.根据权利要求7所述的方法,其特征在于,所述方法包括根据所推荐的组合物来制作和分配化妆品产品的后续步骤。

9.根据权利要求1-8中任一项所述的方法,其特征在于,所述图像中图示的所述区域是所述使用者的发根区。

10.根据权利要求1至9中任一项所述的方法,其特征在于,所述机器学习模型是预训练的卷积神经网络。

11.一种存储有计算机可用指令的非暂态计算机存储介质,所述计算机可用指令在被计算设备使用时使所述计算设备执行根据权利要求1至10中任一项所述的方法。

12.一种用于确定使用者的角蛋白表面的至少一个物理和/或化学特性的系统,所述系统包括图像传感器,所述图像传感器被配置为获取所述角蛋白表面的图像并将图像数据传输至处理器,所述处理器被配置为根据权利要求1至11中任一项所述的方法来处理所述图像数据。

13.根据权利要求12所述的系统,其特征在于,所述系统至少部分地集成到便携式相机设备中。

14.根据权利要求12所述的系统,其特征在于,所述系统包括相机设备,所述相机设备包括所述图像传感器,所述相机设备被配置为将所述图像数据无线地传输到远程处理设备。

15.一种用于制作个性化化妆品产品的系统,所述系统包括:根据权利要求12至14中任一项所述的特性确定系统;以及分配单元,所述分配单元被配置为根据从所述特性确定系统获得的特性来输送化妆品产品。

用于确定角蛋白表面的特性的方法和系统以及用于处理所述角蛋白表面的方法和系统

[0001] 本发明涉及一种用于确定角蛋白表面的至少一个特性的方法和系统,并且涉及一种用于处理所述角蛋白表面的方法和系统。

[0002] 该方法特别涉及人类头发的评估和处理,并且更具体地,涉及一种用于对头发进行漂白或染色的方法和系统,诸如例如文献WO 2004002300 A2和US 9316580 B2中披露的。

[0003] 健康和美容行业利用技术进步来改善消费者对其产品和服务的体验。提供根据使用者需求量身定制并符合其特定特征的产品是深层次的趋势。这种趋势通常称为“个性化”。

[0004] “个性化”可以涉及任何身体部位,但是特别关注的是裸露的身体部位(诸如脸部(化妆产品或护理产品,特别是粉底)和头皮/头发(例如护理产品或染色产品))。

[0005] 已知的是基于使用者的皮肤的至少一个特性来提供和推荐粉底化妆产品(诸如LANCÔME®的“Le Teint Particulier”®)。还已知的是基于头发的某些特性来提供和推荐头发产品。

[0006] “个性化”不仅仅是提供各自与专门的一类使用者相对应的或多或少大范围的产品,挑战通常是能够根据相关使用者的个人特征向其推荐最适当的产品。

[0007] 通常,“个性化”包括旨在获得特定于所述使用者的数据的第一诊断步骤,然后使用该数据来确定要施加于使用者的相关身体部位的(多个)适当的后续处理和/或产品。

[0008] 直到最近,一些或所有步骤经常是由诸如美发师或美容顾问等专家手动地或在视觉上执行的。例如,已知向使用者发出特定的问卷,使用其答案来确定所谓的最合适的产品(例如,US-A1-2014/0216492)。当然,这种方法具有很大的可变性和主观性,并且通常需要同时提高产品推荐的相关性和可靠性。

[0009] 为了提高产品推荐的相关性和可靠性,可以使用工具和设备来执行一些或所有步骤,这些步骤甚至可以是自动化的。

[0010] 例如,可以通过使用适当的测量设备(称为“读取器”)进行测量来从使用者的相关身体部位获得数据,以获得一个或多个特定特征,该一个或多个特定特征被照看以确定适当的处理或产品。这种设备的示例是X-RITE销售的CAPSURE读取器,并且该读取器能够返回使用者皮肤的颜色数据代码。

[0011] 然后,所收集的数据(可能与其他附加数据一起)可以被馈送到计算单元,该计算单元将根据各种规则集来帮助确定最合适的产品。

[0012] 根据所收集的数据确定的后续处理或产品可以是从目录或数据库中选择的现成产品(诸如在披露了用于查找匹配的粉底产品的方法的文献US 5478238或US 9519927中)或者个性化产品,已基于先前获得的数据确定了该个性化产品的组合物。

[0013] 然后,使用者可以购买或订购推荐的产品。在个性化产品(或定制产品)的情况下,然后可以通过能够根据个性化组合物混合组分并分配所述组合物的设备在现场、直接在商店或在美容/美发沙龙中制造所述个性化产品。使用者还可以订购个性化产品,以用于稍后的输送。

[0014] 作为此类系统的完整示例,上述文献US 9316580 B2披露了一种执行对角蛋白纤维的定制处理的方法,该方法包括使用被配置为获取头发光谱的光学读取器进行的诊断步骤,所述光谱用于计算随后的头发处理组合物,然后通过相应的设备制作和分配适当的组合物。也可以参考文献WO 2011024160 A1。

[0015] 如上所述,从使用者那里获得可靠、客观且相关的数据可能特别具有挑战性。对于头发或头皮来说,由于其特定的质地和环境,尤其具有挑战性。

[0016] 取决于寻找的期望信息,可以使用不同的技术:

[0017] 文献KR-A1-101456942披露了一种头发和头皮状况诊断设备,该设备包括传感器单元,该传感器单元用于通过基于由头发或头皮状况诊断出的诊断信号比较先前存储的头发样本或头皮样本来客观地诊断使用者的头发或头皮状况。特别地,该设备可以包括用于获取头皮区域的图像的图像传感器。然后可以通过图像分析模块使用例如光强度识别、颜色识别或对比分析来处理所获取的图像,以便确定头发的某些特性,诸如头发的数量、直径等。尽管该文献指出头发和头皮的对比很强,并因此可以在图像中将它们彼此清晰地区分开,但实际上通常不是这种情况,因为取决于头发所含的黑色素水平,头发至少部分地是半透明的。对于最难识别和计数的稀薄头发、白发和金发尤其如此。

[0018] 以本申请人的名义提交的文献WO 2017207455还提出了一种用于分析头皮和/或头发的状况的方法,该方法包括以下步骤:在不同的照明条件(白光和UV-蓝光)下拍摄图片,以便改善头发的检测,并且更好地展现浅色和白色的头发。然后通过对比分析以互补的方式来处理图像。

[0019] 文献US 2012253203披露了一种用于对呈现头发的皮肤表面(诸如头皮区域)进行成像以确定皮肤和头发生长状况的方法。该方法包括在不同的视角下拍摄相关表面的图片,然后通过计算机分析潜在地处理所述图片,以确定各种特性,诸如从皮肤表面突出的头发的数量、头发的厚度、头发的直径或截面形状、从该表面突出的头发的长度。

[0020] 文献US 2009036800涉及一种头发密度计。密度计通过放大头皮区域的图片来工作。图像中可见的头发可以由技术人员或任何合适的自动化系统来计数和测量。

[0021] 取决于要确定的特定特征,必须开发出特定的算法来处理图像。由于这种算法专用于处理特定特征,因此可能需要较高的计算能力且缺乏通用性。在头发特性测量的特定领域,头发的本质使得设计算法特别困难。

[0022] 另外,在头发染色或漂白的特定领域,对原始头发颜色进行可靠且真实的诊断至关重要,因为它将对要获得的最终颜色产生强烈影响。通常,这是由专家(诸如使用者的美发师)根据其专业知识以及在示出不同头发色调和头发颜色的标度的帮助下在视觉上完成的。

[0023] 2003年初,现在隶属于欧莱雅(L'OREAL)集团的COLORIGHT公司提出了一种用于基于使用者的头发的颜色来确定和制备头发染料组合物的系统。颜色数据通常是通过使用分光光度计获得的,该分光光度计的光谱被用于计算适当的头发染料组合物并预测最终的染色结果。有关COLORIGHT的应用的附加信息,可以参考先前所述文献WO 2004002300和后续申请。随后,先前所述的专利US 9316580 B2也与针对全球系统的描述特别相关。

[0024] 存在涉及测量使用者的头发颜色的若干文献。

[0025] 可以参考文献US 6067504,其披露了一种用于使用色度计来正确识别头发颜色的

方法。该方法进一步使用头发的Hunter L、a、b值,并使用这些值来从预定的数据库中提出适当的着色剂,该着色剂可以实现期望的头发色调。

[0026] 使用分光光度计测量头发颜色的一个特殊问题是,它采集整个测量区域的光谱,所述区域包括头皮皮肤和头发。分光光度计能够高准确度地测量对象的颜色,但是它们会在整个测量区域(通常为1cm^{2}的数量级)上对颜色的值取平均值。因此,实际上从光谱中获得的颜色并不完全对应于使用者头发的真实颜色,并且可能需要校正。

[0027] 在头发颜色测量的特定背景中,不能在头皮附近使用分光光度计进行根部测量,因为分光光度计会对整个区域求平均,从而将头发值和头皮值混合在一起,并且高度依赖于被测量的确切区域。

[0028] 此外,分光光度计或色度计专用于颜色测量,并且不可能获得附加的相关数据(诸如头发的密度等),或者需要附加的传感器或附加的专用设备才能获得这些数据。

[0029] 为了提高测量的可靠性,文献US 7508508披露了光谱分析单元和图像单元的组合用途,所述图像单元用于显示被测量区域的图像,从而允许操作员精确地控制被测量区域并执行互补的视觉评估。这种解决方案仍然很复杂并且不是完全令人满意的。

[0030] 也可以参考文献US 7151851,该文献描述了一种化妆品颜色分析系统,该系统通过相机传感器的采集来分析三维对象(诸如头发)的颜色,并通过对多个化妆品颜色进行相对加权来确定化妆品颜色。然而,该系统没有考虑头皮和头发的环境。

[0031] 在另一种方法中,文献US 8428382提出了一种用于显示头发区的图像以便使用图像处理来评估头发造型的方法和装置。

[0032] 然而,尽管进行了各种试验,但是成功地转换头发诊断的功能设备仍然是一个具有挑战性的课题。

[0033] 需要进一步开发该系统并允许多功能地、可靠且客观地收集头皮数据和头发数据。另外,头发颜色和头发密度不是唯一的相关参数,并且可能希望收集特定于使用者的其他参数。

[0034] 为此,需要一种用于诊断皮肤、头皮和/或头发的状况的改进方法,以便获得特定于使用者的可靠数据,然后可以将该数据用于个性化产品推荐系统或用于个性化的分配或组合物制造系统。

[0035] 本申请提出了一种针对上述限制中的至少一部分限制的解决方案,并且涉及一种用于确定使用者的角蛋白表面的至少一个物理和/或化学特性的方法,该方法包括以下步骤:

[0036] -接收与该角蛋白表面的至少一个图像相对应的数据,(输入图像)

[0037] -通过对所述图像应用至少一个机器学习模型来处理该图像,

[0038] -返回与该角蛋白表面的要确定的特性的评分等级相对应的至少一个数值。

[0039] 确实已经出乎意料地发现,在利用适当的训练集进行训练之后,应用于使用者皮肤图像的机器学习模型可以在评估和分级要测量的角蛋白表面的物理和/或化学特性方面提供良好的结果。

[0040] 由于它是基于图像的(机器视觉)并且不需要特定的传感器,因此易于使用和实施,同时还发现通过使用机器学习模型分析图像可以帮助评估那些肉眼(甚至是专家的眼睛)很难或根本无法评估的特性。

[0041] 通过返回或输出(与通过分类模型预测的标签不同的)数值,所提出的方法允许定量地确定期望的特性。然后可以将定量值用作进一步计算和转换中的输入数据。这不仅仅是其中返回种类或类别的分类,即使类别的标签可能使用数字也是如此。

[0042] 更确切地说,返回的数值是连续的输出变量,它是诸如整数或浮点值等实值。这些值通常是量,诸如数量和大小。

[0043] 机器学习模型是回归模型。可替代地,机器学习模型是其中标签具有排序关系的分类模型(离散化模型),所述标签被排序为与要估计的物理/化学特性相对应的连续范围。

[0044] 重要的是要注意,尽管分类模型可以预测连续值,但是所述值采用种类标签的概率的形式。回归模型可以预测离散值,但是所述离散值是整型量的形式。

[0045] 优选地,图像是高分辨率图像。

[0046] 优选地,图像是角蛋白表面的低倍放大图像。如通常所理解的,低倍放大图像是以1:1或更高的放大率拍摄的图像。优选地,每个像素的最大大小为10μm,从而沿着发丝的直径具有若干个像素。

[0047] 可以通过诸如CMOS或CCD图像传感器等已知图像传感器获得图像数据。另外,图像有利地是无视差的。

[0048] 有利地,该方法包括预处理图像分割步骤,即,将机器学习模型应用于至少一个识别出的分段。图像分割旨在提取与要确定的特性相关联的重要对象。图像分割是将数字图像分为多个分段或像素集以将图像表示简化和/或更改为更有意义且更易于分析的内容的过程。图像分割通常用于定位图像中的对象和边界(线条、曲线等)。这种预处理步骤将通过突出显示图像数据中的特定对象来增强由机器学习模型做出的预测。例如,可以对头皮的图像进行分割,以识别出很可能是头发的像素。然后可以将机器学习模型应用于这些特定的分段,以进行更好的预测。

[0049] 根据第一实施例,分割步骤通过对比分析来执行。根据第二实施例,通过将机器学习模型(即分类模型)应用于图像来执行分割步骤。

[0050] 可以对图像数据执行其他预处理步骤,诸如颜色校准和校正。

[0051] 优选地,图像是RGB图像。可替代地,图像可以是CMYK图像,或者可以在另一颜色空间中编码。可替代地,图像是多光谱图像,该多光谱图像包括选自380nm至700nm的可见光谱带、300nm至380nm的UV带(或者甚至是200nm至380nm的近UV带)和700nm至1500nm的IR带的至少一个光谱带。图像也可以是红外图像。

[0052] 在特定的实施例中,重复该方法以确定角蛋白表面的第二特性,优选地使用相同的图像数据重复该方法。实际上,通过简单地加载适当的模型,可以很容易地对同一图像进行处理以获得另一个特性的估计值。优选地,即使当应用不同的机器学习模型时(这些模型被应用于同一单个图像),该方法也使用单个图像的数据。

[0053] 有利地,该方法包括基于至少一些实值特性来推荐化妆品产品的后续步骤。众所周知,可以根据其列出的特性在数据库中选择该产品。

[0054] 根据本申请,“化妆品产品”或“化妆品组合物”应理解为是指如2009年11月30日欧洲理事会和议会法规1223/2009中定义的属于化妆品产品的产品。

[0055] 在特定的实施例中,该方法包括基于至少一些实值特性来确定推荐的化妆品产品组合物的后续步骤。

[0056] 当然,在确定推荐的产品和组合物时,可以综合考虑通过不同方式(测量、问卷调查、图像分析)获得的附加参数和特性。

[0057] 有利地,该方法包括根据所推荐的组合物来制作和分配化妆品产品的后续步骤。特别地,该方法包括在混合单元中混合至少一些组分的步骤。在头发染料的特定情况下,该方法包括将至少一种头发染料着色剂(优选地以固体形式,如珠子或片剂)分配到头发染料氧化显色碱剂(优选地以液体或乳膏的形式)中的步骤。

[0058] 优选地,图像中图示的角蛋白表面是使用者的皮肤区域。在第一实施例中,图像中图示的皮肤区域是使用者的脸部区域、优选地是使用者的脸颊和/或前额的区域。可替代地,图像中图示的皮肤区域是使用者的头皮区域。在最优选的实施例中,图像中图示的皮肤区域包括头发。

[0059] 有利地,图像中图示的区域是使用者的发根区。头发的根部(即距头皮的第一厘米)为我们呈现了干净的头发纤维部分,这些部分未因头发染色或环境状况而发生颜色变化。因此,它们是人的基本头发特性的度量。

[0060] 根据本申请,要确定的特性选自:头发色调、头发直径、头发密度、白发百分比、长度颜色、根部颜色、头发光泽、头皮干燥度、头皮屑程度、头发比色、头发光谱、真黑色素水平、类黑色素水平、人造染料水平、湿度、磺基丙氨酸水平、损伤水平、竖起水平,已经对机器学习模型进行了相应的训练以评估所述特性。

[0061] 优选地,机器学习模型是预训练的卷积神经网络。

[0062] 在第一特定实施例中,该方法涉及确定使用者的自然头发色调,该方法包括以下步骤:

[0063] -接收与使用者的头皮表面的图像相对应的数据,所述图像包括优选地沿着图像的中位线分开的头发,

[0064] -对图像进行分割,以提取与可能的头发相对应的像素集,

[0065] -将分割后的头发像素的颜色空间值转换为CIELAB颜色空间,

[0066] -基于分割后的头发像素的中位亮度值返回实值的头发色调估计结果。

[0067] 在第二特定实施例中,该方法涉及确定使用者的自然头发色调,该方法包括以下步骤:

[0068] -接收与使用者的头皮表面的图像相对应的数据,所述图像包括优选地沿着图像的中位线分开的头发,

[0069] -通过应用预训练的卷积神经网络返回实值的头发色调估计结果。尽管有可能,但该方法不需要分割步骤,因为看起来经过适当预训练的CNN不需要分割即可提供良好的结果。

[0070] 本申请还涉及一种或多种存储有计算机可用指令的非暂态计算机存储介质,这些计算机可用指令在被计算设备使用时使该计算设备执行根据本发明的方法。

[0071] 本申请还涉及一种用于实施根据本发明的方法的系统,并且更确切地涉及一种用于确定使用者的角蛋白表面的至少一个物理和/或化学特性的系统,所述系统包括图像传感器,该图像传感器被配置为获取该角蛋白表面的图像并将图像数据传输至处理器,所述处理器被配置为根据本发明的方法来处理该图像数据。

[0072] 如上所述,图像传感器可以是本领域已知的传感器,诸如CCD传感器或CMOS传感

器。取决于期望的图像,传感器也可以是多光谱传感器和/或使用诸如红外滤光器等滤光器以获得红外图像。

[0073] 有利地,该系统至少部分地集成到便携式相机设备中。

[0074] 优选地,该系统包括相机设备,该相机设备包括图像传感器,所述相机设备被配置为将图像数据无线地传输到远程处理设备。

[0075] 可替代地,相机设备还可以包括处理单元、并且直接输出返回的特性值。通过这种方式,相机设备形成了可以易于手持并用作独立设备的自主诊断单元。

[0076] 本申请还涉及一种用于制作个性化化妆品产品、特别是头发护理产品或头发着色产品的系统,该系统包括根据本发明的特性确定系统(优选地,单元,即手持式的)以及分配单元,该分配单元被配置为根据从该特性确定系统获得的特性来输送化妆品产品。有利地,特性确定系统可以将实值数据无线地传输到分配单元。更确切地说,特性确定单元会将实值特性作为输入转发给计算单元(推荐),该计算单元将确定合适的输出产品或组合物。推荐单元然后可以将输入指令发送到混合和输送单元,以便向使用者提供推荐的产品或组合物。

[0077] 根据参照附图进行的以下详细说明,将更好地理解本申请的主题,在附图中:

[0078] 图1总体上展示了根据本发明的用于确定使用者的自然头发色调的各种顺序步骤。

[0079] 图2总体上展示了获得要应用的机器学习模型时涉及的训练过程。

[0080] 图3示出了使用比色模型的步骤流水线。

[0081] 图4展示了应用于所获取的头皮图像的各个层。

[0082] 尽管本申请与头皮和头发诊断以及头发特性测量特别相关,但是本申请不限于这种角蛋白表面,并且可以在更一般的皮肤表面表征和测量中进行应用。

[0083] 如上所述,在这个数字化时代,头发护理和头发着色是面临许多挑战的行业。由于头发的复杂性和头发染色的过程,准确的头发诊断对于为客户提供个性化的头发护理产品和头发着色产品至关重要。

[0084] 以可靠的方式获得头发相关的特性是最具挑战性的,下面的描述通过头发测量来说明。然而,如上所述,本方法不限于头发测量并且可以进行其他皮肤应用。

[0085] 当前,在美发沙龙中,在应用任何头发着色产品或建议进行头发处理之前的第一个必要步骤是进行头发诊断。美发师需要估计的重要特征包括头发色调、白发百分比、头发直径和密度。

[0086] 头发的根部通常是我们能够接触到未被外界因素(诸如天然/人造染色剂)改变的头发的唯一区。在这个区中,我们可以测量自然的头发颜色和白发百分比以及头发密度。

[0087] 头发的根部(即距头皮的第一厘米)为我们呈现了干净的头发纤维部分,这些部分未因头发染色或环境状况而发生颜色变化。因此,它们是人的基本头发特性的度量。

[0088] 在根部区获取的图像不仅示出头发,而且还示出头皮,该头皮的颜色、油性和头皮屑含量可能变化很大。另外,头发纤维本身、特别是在根部部分是半透明的(从而导致颜色依赖于头皮背景),并且在颜色和厚度上都具有一定的自然变化性。当前,这种诊断是由美发师手动进行的,并且尽管他们具有专业知识并受过训练,但他们并不总是能够对所有这些特征进行准确的估计,尤其是在美发沙龙的非标准照明条件下。

[0089] 在根部建立准确的头发诊断方法是一项重大挑战,对头发着色、美容个性化和临床评估都有巨大影响。

[0090] 将针对头发色调测量(尤其是自然头发色调测量)详细地说明根据本申请的方法,但是该方法可以更普遍地用于确定其他物理或化学特性,这取决于模型训练所针对的特性。

[0091] 自然头发色调是头发亮度/暗度的视觉属性,并且与头发黑色素浓度有关。因此,它受到自然界中存在的头发颜色的颜色空间的约束。传统上,它是使用对数标度来衡量的,从1(对于非常深色的头发)开始到10(对于非常浅的金发)。

[0092] 由于将头发色调分为10个类别,因此可以将这一问题作为分类问题来解决。然而,色彩专家进行色调评估的精度为1/4个头发色调,因此考虑1到10个种类将在种类标签上失去这种准确性。因此,已经将该问题作为回归问题来解决,它假设了头发色调值之间的顺序和连续性。因此,该方法估计并返回实值的头发色调。

[0093] 关于用于评价头发特性的本方法,重要的是,头发色调标度是感知线性标度表示法,其适于与本方法一起使用以返回数值(预测)而不是返回种类本身(分类)。

[0094] 根据本申请,确定使用者的自然头发色调包括以下步骤,这些步骤的一般顺序如图1所示。

[0095] 首先,使用形成诊断单元的相机设备20拍摄使用者的头皮10在头发根部11处的图像I。相机设备20优选地是手持式的。图像I对应于约12×17mm的对象(头皮)区域。图像I是通过2000×1200RGB CMOS传感器获取的,但也可以使用其他种类的高分辨率传感器。

[0096] 在获取之前,在受试者的头部上分出一条线,以使头发根部11上的视野不被遮挡。因此,图片围绕其中间具有轴对称性,并且通常被这样定向,其中头皮10在中间,并且更多的头发在顶部和底部。

[0097] 该方法用于估计无偏头发色调值h。然而,从所获取的图像I不能直接读取该头发色调h,因为这些图像呈现出头发11与头皮10之间的混合。而且,由于头发透明度,头皮的颜色会影响可见的头发像素。

[0098] 因此,应当以获得准确的头发色调h的特定目的对图像I数据进行处理。根据本发明,通过在返回数字实值头发色调之前将机器学习模型应用于所述图像来处理图像数据。

[0099] 下文描述的机器学习模型已使用包含对407个受试者拍摄的11175个图片的训练集进行了统计训练。每个图片都经过评估,并用期望属性的实际值标记(训练过程总体上如图2所示)。

[0100] 头发色调估计使用了以下两种方法,即比色模型和基于卷积神经网络(CNN)的深度学习模型。

[0101] 比色模型

[0102] 该方法背后的思想是将成像设备转换为比色设备,该比色设备能够提供头发的标准颜色值并将标准颜色值与头发色调的感知属性进行相关。将这种方法的完整流水线分为图3所示的以下三个步骤。

[0103] 首先,通过分割步骤S对获取的图像I进行预处理。更确确地说,从图像中分割出感兴趣的头发像素。然后,将分割后的头发像素的RGB值转换为CIELAB颜色空间,并且拟合的比色模型M根据头发的中位L*值来估计头发色调。因此,这种比色方法涉及使用依次应用于

图像数据的两个独立的机器学习模型:首先,应用机器学习模型以进行颜色空间变换,然后,应用第二个机器学习模型以根据中位头发L*进行头发色调估计。

[0104] 头发分割S是第一预处理步骤。此步骤的主要目的是将头发像素的不同RGB值转换为单个整数L*a*b*值。因此,代替对图像中的所有头发像素进行分割(这可能会引入离群值),优选的是对足够数量的头发像素进行鲁棒地分割,以避免出现镜面反射和其他伪影。分割方法基于自适应阈值化,以便极大地对头皮区域的头发进行分割。在阈值化之前,应用高斯滤波器以减少图像中的噪声。最后,使用连接的组分分析按大小对产生的头发区进行排序,以消除检测到的、可能是镜面反射或被错误地分割为头发的其他伪影的较小区域。

[0105] 然后,通过应用机器学习模型将RGB值转变为CIELab颜色空间。更具体地,应当对相机设备进行校准,以便将与设备相关的RGB值转换为与设备无关的颜色空间。选择CIELAB颜色空间进行此转换是因为它在感知上相对统一,并且在行业中被广泛使用。

[0106] 文献中有各种方法可提供对头发的L*a*b*测量。对于校准过程,建议使用以下模型来估计表示为$r \in \mathbb{R}^3$的每个RGB像素的L*a*b*值$r \in \mathbb{R}^3$:

[0107] $$\hat{r}(x) = C_r \phi(x)$$

[0108] 其中,φ(x)是RGB像素值x的三次多项式(N=20),并且C_r是对应的系数,此系数是通过机器学习模型根据一组K个训练对统计学习得到的。这些对是通过获取具有测量的L*a*b*值r_i的预定义色块的图像来获得的。对于设备获取的每个块x_i,我们取中位RGB像素以求解回归问题:

[0109] $$\arg\min_{C_r} \sum_{i=1}^{K} \|\hat{r}(x_i) - r_i\|^2$$

[0110] 这是C_r的闭型方程并且$||\cdot||^2$表示L^2范数。

[0111] 一旦分割出感兴趣的头发像素,就将其RGB值转换为CIELAB空间,并保持中位L*a*b*值。已经注意到,在不同的头发色调之间变化最显著的感知值是亮度。为了保持模型简单并避免数据过度拟合,选择L*值作为与头发色调最相关的值。使用训练集,在M对分别表示为l_i和h_i的L*与头发色调值之间拟合得到一维三次平滑样条$\hat{f}(l)$。这些L*和头发色调值对是通过对一个参与者的所有对应图片求平均来计算的。我们自动选择样条节点的数量作为最小数量,从而满足以下条件:

[0112] $$\frac{1}{M}\sum_{i=1}^{M} \|h(x_i) - \hat{f}(l_i)\|^2 \leq \sigma^2$$

[0113] 平滑因子σ被设置为0.35,以便忽略沿头发色调范围[1;10]的细微变化。

[0114] 总而言之,根据单个输入图像I预测头发色调。我们分割出I中的鲁棒的头发像素x的子集。然后,在使用C_rφ(x)将各像素x变换为L*a*b*空间之后,我们计算这些像素的中位L*值l(I)。最后,通过应用机器学习模型来给出头发色调:

[0115] $$\hat{h}(I) = \hat{f}(l(I))$$

[0116] 卷积神经网络

[0117] 这种方法背后的思想是学习图像中的哪些图案与头发色调有关,而无需事先假设要提取的图案。与设计用于关注头发像素亮度的先前比色模型相反,该第二个模型没有针

对该问题的先验知识。优化过程纯粹是统计过程,并且会自主学习图像I中的图案,以估计正确的头发色调$\hat{h}(I)$。

[0118] CNN背后的基本思想是应用连续的线性运算(诸如3×3卷积)以及它们之间的非线性函数、以及降低图像表示的空间维度的运算。这将产生图像的特征图,该特征图是视觉图案的表示。在逐步回归之后,这种表示被简化为单个实值输出,该输出最终将在训练过程中映射到标签h_i。

[0119] 所有这些运算都表示为如图4所示的、我们针对本头发色调应用提出的网络的连续的层。与现有技术的CNN模型相比,它的卷积层更少;这是因为我们正在寻找更简单的图案(诸如边缘和细纤维),从而需要由网络表示的连续卷积较少。

[0120] 最后的密集层也减少了,因为我们具有一维输出$\hat{h}(I)$。而且.为了便于优化(反向传播中的直接链接)并利用更简单的模式扩展模型表示,向架构中添加了残差连接。

[0121] 为此,我们将一些层的输出添加到了后一层,这是不需要额外权重的轻量级运算。为了提高速度,我们使用可分离的2D卷积,作为常规2D卷积的更轻量级替代方案。它们需要学习的参数较少,并且所需的计算也较少。这个思想是首先逐通道(也称为深度方向)地应用3×3卷积(这样比全卷积更轻量级),并且然后通过1×1全卷积来组合得到的通道。卷积的这种变体将加快训练过程以及设备上的嵌入式预测。

[0122] 现在,我们简要解释这种模型的优化过程,以便对神经网络如何学习我们问题的相关模式提供一些直观认识。如果我们从连续运算中退后一步,我们会看到我们估计的头发色调可以写成

[0123] $$\hat{h}(I) = g(I,\Theta)$$

[0124] 其中,g表示我们神经网络的所有连续运算,并且Θ是将所有变量参数(即,卷积权重和密集层权重)重新分组的向量。为了找到Θ的适当值,我们定义损失函数$\mathcal{L}\left(\hat{h}(I),h\right)$,该损失函数对远离地面真值h的预测$\hat{h}(I)$进行罚分。在本文的情况下,我们最小化关于我们数据集的所有图像I_i和头发色调h_i的均方误差$\mathcal{L}\left(\hat{h}(I),h\right) = (\hat{h}(I)-h)^2$:

[0125] $$\min_{\Theta}\sum_{i=1}^{M}\mathcal{L}\left(\hat{h}(I_i),h_i\right) = \min_{\Theta}\sum_{i=1}^{M}(g(I,\Theta)-h_i)^2$$

[0126] 我们解决此最小化问题的策略是使用随机梯度下降法来迭代更新Θ:

[0127] $$\Theta' = \Theta + \epsilon\widehat{\nabla_{\Theta}\mathcal{L}}$$

[0128] 其中,∈是学习速率,并且$\nabla_{\Theta}\mathcal{L}$是整个数据集上的梯度的近似值。实际上,对于每次更新,这个完整的梯度$\sum_i \nabla_{\Theta}\mathcal{L}\left(\hat{h}(I_i),h_i\right)$都过长而无法计算。使用反向传播,可以针对大小为m个图像的小批量(minibatch)高效地计算梯度。Θ因此通过考虑如下梯度的移动平均值在每个小批量中进行更新

[0129] $$\widehat{\nabla_{\Theta}\mathcal{L}'} = \lambda\widehat{\nabla_{\Theta}\mathcal{L}} + (1-\lambda)\frac{1}{m}\sum_{i\in\mathcal{B}}\nabla_{\Theta}\mathcal{L}\left(\hat{h}(I_i),h_i\right)$$

[0130] 其中,λ是动量,并且B是小批量的索引i的集合——$\mathcal{B}\subset\{1..M\}\ |\mathcal{B}| = m$。

[0131] 因此,权重的每个小批量更新均基于梯度的历史,而不仅是基于小批量梯度$\sum_{i\in\mathcal{B}}\nabla_{\Theta}\mathcal{L}\left(\hat{h}(I_i),h_i\right)$,从而通过小批量呈现很大的变化。在实践中,我们对我们数据集的所有样本进行了800次通过,其中学习速率为$\in=10^{-4}$并使用$\lambda=0:9$的动量。最后,为了进一步调整参数Θ,我们对我们的数据集又进行了20次通过,其中学习速率除以10。在每次通过期间,每个图像Ii都会被随机地水平翻转和竖直翻转,以便人为地增强我们的数据集。对于我们的连续层和优化过程,我们使用了基于Tensorflow后端的Keras实施方式。

[0132] 尽管通过两个机器学习模型(比色模型和卷积神经网络模型)进行了说明,但也可以使用其他机器学习技术,本申请的主要方面是使用统计机器学习模型代替经典计量学,从而使用图像分析技术直接提取相关参数并且实际执行测量。

[0133] 然后,返回的实值特性可以用作进一步计算或步骤的参数。

[0134] 作为第一种可能性,可以在推荐器中使用实值特性,该方法包括基于所述实值特性来推荐化妆品产品的后续步骤。

[0135] 作为第二种可能性,代替推荐现有产品,可以使用实值特性基于期望的效果或结果来计算或确定化妆品产品的组合物。更具体地,该方法可以包括以下步骤:基于实值特性的变化能力来计算组分的比例和/或量,以便实现期望的结果。例如,如果使用者希望对其头发进行染色以达到5的头发色调并且其初始头发色调为7,则该方法将包括计算达到该结果所需的氧化剂量的步骤。

[0136] 然后可以在现场或非现场制造所确定的组合物,并相应地将其分配或输送给使用者。

[0137] 应当理解,本发明的各方面可以具体化为一种系统、方法或计算机程序产品。因此,本发明的各方面可以采取完全硬件实施例、完全软件实施例(包括固件、驻留软件、微代码等)或结合软件和硬件方面的实施例的形式。此外,本发明的各方面还可以采取在一个或多个计算机可读介质中具体化的计算机程序产品的形式,该一个或多个计算机可读介质具有在其上具体化的计算机可读程序代码。

[0138] 可以利用一个或多个计算机可读介质的任何组合。计算机可读介质可以是例如但不限于电子的、磁性的、光学的、电磁的、红外的、或半导体系统、装置、或设备、或前述各项的任何合适的组合。计算机可读存储介质的更具体的示例(非详尽列表)将包括以下各项:具有一根或多根导线的电连接、便携式计算机磁盘、硬盘、固态硬盘、随机存取存储器(RAM)、只读存储器(ROM)、可擦除可编程只读存储器(EPROM或闪速存储器)、光纤、便携式致密盘只读存储器(CD-ROM)、光学存储设备、磁存储设备、相变存储器存储设备或前述各项的任何合适的组合。在本文献的上下文中,计算机可读存储介质可以是可包含或存储用于由指令执行系统、装置或设备使用或与指令执行系统、装置或设备结合使用的程序的任何有形介质。

[0139] 用于执行针对本发明的各方面的操作的计算机程序代码可以按一种或多种编程语言的任何组合来编写,这些编程语言例如面向对象编程语言(诸如,Java、Smalltalk、C++等)或常规程序编程语言(诸如,“C”编程语言或类似的编程语言)。程序代码可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为独立软件包执行、部分地在用户计算机上执行并部分地在远程计算机上执行、或完全地在远程计算机或服务器上执行。在后一种情况下,远程计算机可以通过包括局域网(LAN)或广域网(WAN)的任何类型网络连接到用

户计算机,或可以进行与外部计算机的连接(例如,使用互联网服务提供商通过互联网)。应当理解,计算机领域中的普通技术人员可以基于说明书中所包含的功能说明和附图中所示的流程图来开发本发明实施例的计算机系统的软件。进一步,本文中对执行各种功能的软件的任何提及通常涉及在软件控制下执行那些功能的计算机系统或处理器。

[0140] 计算机系统的各种功能可以以任何方式分布在任何数量的软件模块或单元、处理系统或计算机系统和/或电路系统中,其中,这些计算机系统或处理系统可以本地或彼此远程地设置,并经由任何合适的通信介质(例如,LAN、WAN、内联网、互联网、硬线、调制解调器连接、无线等)进行通信。

[0141] 更确切地说,以上详细描述的方法被实施到用于确定使用者的角蛋白表面的至少一个物理和/或化学特性的系统中,所述系统包括图像传感器,该图像传感器被配置为获取目标角蛋白表面的图像,该图像数据被传输到被配置为根据上述方法(即通过应用机器学习模型)来处理该图像数据并返回特性的实值估计结果的处理器。这种系统形成诊断单元。

[0142] 该系统被集成到便携式相机设备中。在第一实施例中,相机设备包括图像传感器,并且图像数据在相机设备外部、在远程处理单元中被处理。有利地,图像数据是以无线方式传输的。在替代实施例中,该系统完全集成到便携式设备中,这意味着相机设备还包括处理单元。通过这种方式,相机设备可以用作直接输出实值特性的自主和/或实时诊断单元。

[0143] 该系统可以包括附加部件,以便例如形成用于制作个性化化妆品产品(特别是美发产品,即头发着色产品)的系统。为此,该系统包括如上所述的诊断单元以及分配单元,该分配单元被配置为根据从该诊断单元获得的特性来输送化妆品产品。对于分配单元的详细说明,可以参考先前所述的文献US 9316580,其内容完全整合在此。

[0144] 因此,本申请中描述的系统和方法改进了个性化过程的整体自动化。对于头发处理(特别是头发染色处理),该过程的自动化对于方便美发师的生活以及为他们配备能够确保精度、鲁棒性和效率的工具是至关重要的。而且,使用这种设备,我们可以超越人类的视觉感知,并提供不会因美发师和条件的不同而不同的客观表示法,换句话说,就是标准化表示法。

[0145] 前述示例说明了本发明的实施例的某些功能,并不旨在进行限制。实际上,在阅读了本披露之后,其他功能和其他可能的用例对于本领域技术人员而言将是显而易见的。

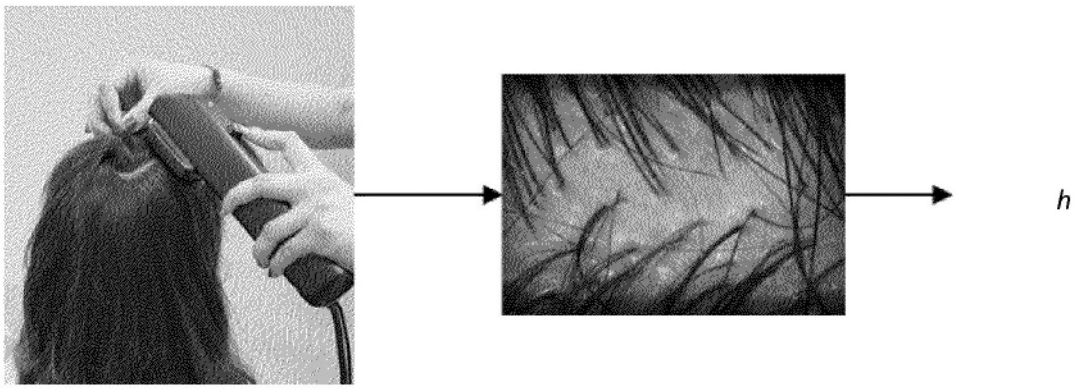


图 1

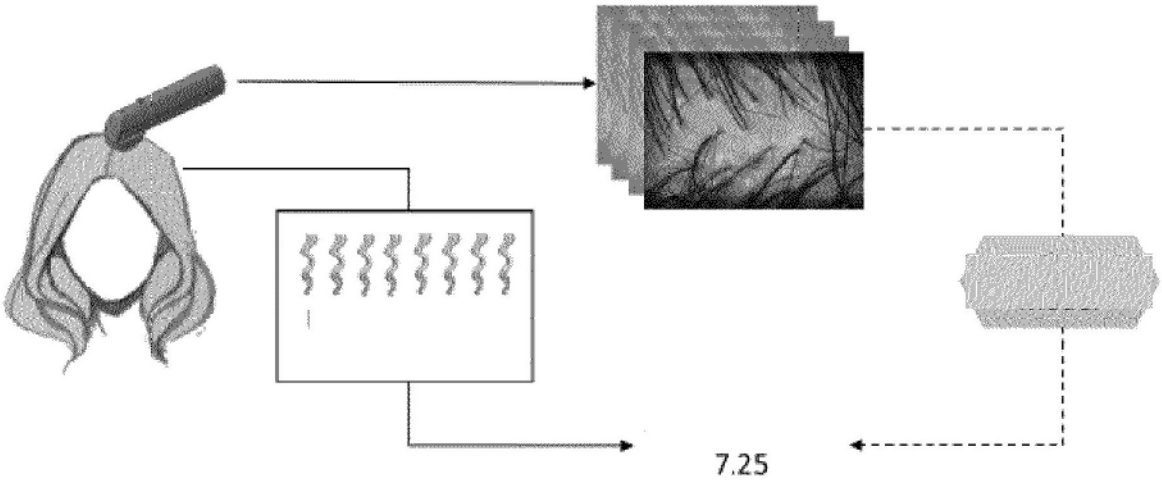


图 2

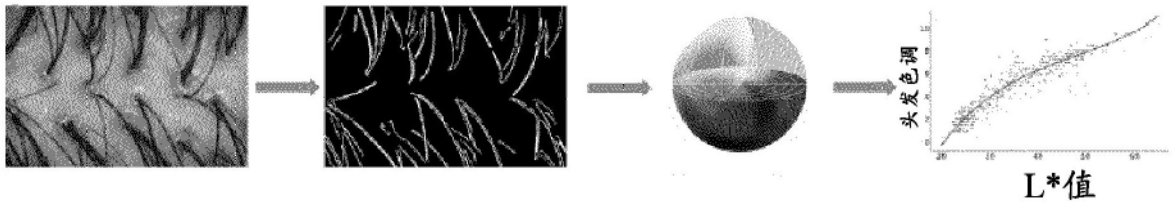
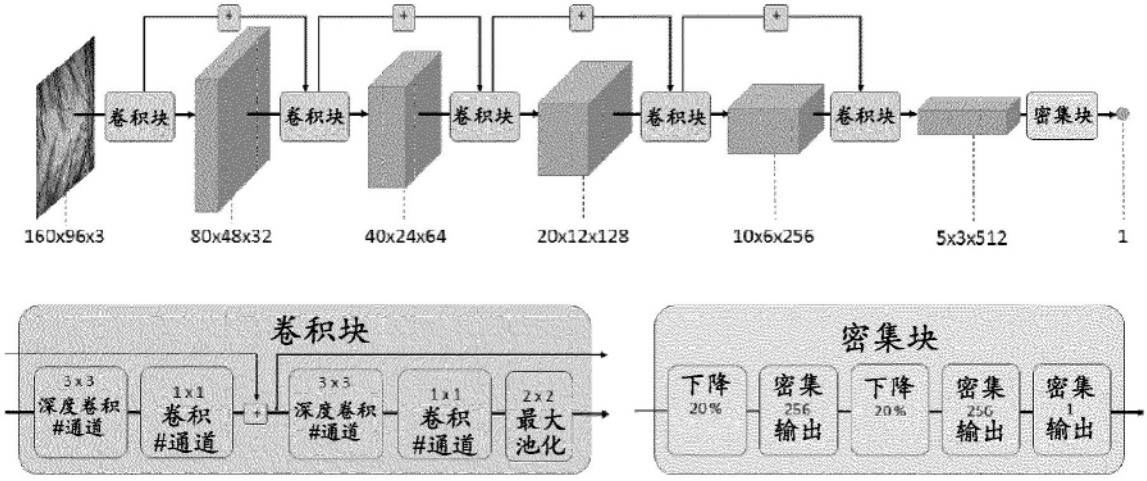


图 3


卷积块
卷积块
卷积块
卷积块
卷积块
密集块
160x96x3
80x48x32
40x24x64
20x12x128
10x6x256
5x3x512
1
卷积块
3x3
深度卷积
#通道
1x1
卷积
#通道
3x3
深度卷积
#通道
1x1
卷积
#通道
2x2
最大
池化
密集块
下降
20%
密集
256
输出
下降
20%
密集
256
输出
密集
1
输出

]

图 4