



(12) 发明专利

(10) 授权公告号 CN 109040756 B

(45) 授权公告日 2021.01.15

(21) 申请号 201810710378.9

H04N 19/172 (2014.01)

(22) 申请日 2018.07.02

H04N 19/51 (2014.01)

H04N 19/56 (2014.01)

(65) 同一申请的已公布的文献号

申请公布号 CN 109040756 A

(56) 对比文件

WO 2016024142 A1, 2016.02.18

WO 2015176678 A1, 2015.11.26

CN 103873861 A, 2014.06.18

(43) 申请公布日 2018.12.18

(73) 专利权人 广东工业大学

地址 510006 广东省广州市越秀区东风东
路729号

审查员 高宇腾

(72) 发明人 刘倩 刘立程

(74) 专利代理机构 广州粤高专利商标代理有限

公司 44102

代理人 林丽明

(51) Int. Cl.

H04N 19/122 (2014.01)

H04N 19/137 (2014.01)

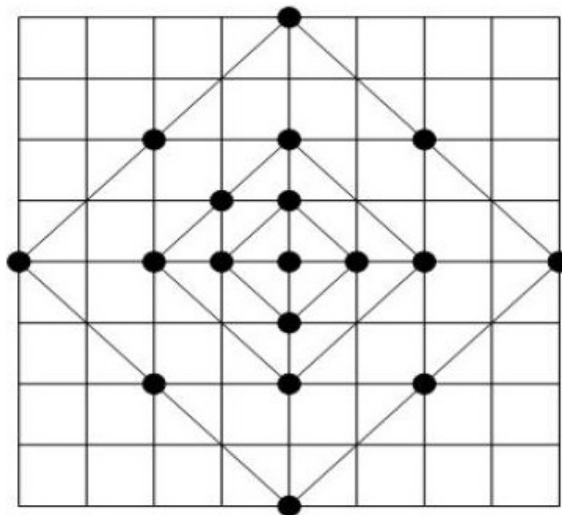
权利要求书1页 说明书5页 附图2页

(54) 发明名称

一种基于HEVC图像内容复杂度的快速运动估计方法

(57) 摘要

本发明提供一种基于HEVC图像内容复杂度的快速运动估计方法。该方法包括如下步骤：S1. 用Merge模式编码当前CU块；S2. 判断当前CU块的运动复杂度，若 $\max\{|d_1-d_2|, |d_2-d_3|, |d_3-d_1|\} < \theta_{th}$ ，且 $\max\{|\theta_1-\theta_2|, |\theta_2-\theta_3|, |\theta_3-\theta_1|\} < \theta_{th}$ ，则进行步骤S3，反之，则进行步骤S4；S3. 用Inter_2N×2N模式编码当前CU块，且只进行步长为2的菱形搜索；S4. 用Inter_2N×2N模式编码当前CU块，进行完整的Test Zero (TZ) Search，并判断最优搜索点是否在初始搜索点附近，若在，进行步骤S5，若不在，则进行步骤S6；S5. 用其他模式（除Inter_2N×2N外）编码当前CU块，进行步长为2的菱形搜索，并确定子PU模式下的最优MV。S6. 用其他模式编码当前CU块，进行完整的TZ Search，并确定子PU模式下的最优MV。本发明能够有效的改进HM中原有TZ Search的编码效率。



CN 109040756 B

1. 一种基于HEVC图像内容复杂度的快速运动估计方法,其特征在于,包括如下步骤:

S1. 用Merge模式编码当前CU块;

S2. 判断当前CU块的运动复杂度,若 $\max\{|d_1-d_2|, |d_2-d_3|, |d_3-d_1|\} < d_{th}$, 且 $\max\{|\theta_1-\theta_2|, |\theta_2-\theta_3|, |\theta_3-\theta_1|\} < \theta_{th}$, 则进行步骤S3, 反之, 则进行步骤S4;

其中, d_1, d_2, d_3 分别为当前CU块到与其相邻的左、上、右上三个方向上已编码CU块的MV大小; $\theta_1, \theta_2, \theta_3$ 分别为 d_1, d_2, d_3 与水平线的夹角; d_{th} 和 θ_{th} 为设定的阈值;

S3. 用父PU模式编码当前CU块, 且只进行步长为2的菱形搜索;

S4. 用父PU模式编码当前CU块, 进行完整的TZ Search, 并判断最优搜索点是否在初始搜索点附近, 若在, 进行步骤S5, 若不在, 则进行步骤S6;

S5. 用子PU模式编码当前CU块, 进行步长为2的菱形搜索, 并确定子PU模式下的最优MV;

S6. 用子PU模式编码当前CU块, 进行完整的TZ Search, 并确定子PU模式下的最优MV;

所述步骤S3和步骤S4中, 父PU模式为Inter_2N×2N模式;

所述步骤S5和步骤S6中, 子PU模式为Inter_2N×N, Inter_N×2N, Inter_N×N, Inter_2N×nU, Inter_2N×nD, Inter_nL×2N以及Inter_nR×2N模式。

一种基于HEVC图像内容复杂度的快速运动估计方法

技术领域

[0001] 本发明涉及高效视频编码 (High Efficient Video Coding, HEVC) 技术领域, 更具体地, 涉及一种基于HEVC图像内容复杂度的快速运动估计方法。

背景技术

[0002] 视觉是人类感知外界的主要途径, 在这个信息化的时代, 视频信息总是会收到人们的青睐。然而未经压缩的视频信息数据量十分惊人, 根本无法用于实际的传输和存储。国际电信联盟电信标准化部门 (International Telecommunication Union-Telecommunication Standardization Sector, ITU-T) 与国际标准化组织 (International Electrotechnical Commission, IEC) 相继推出了一系列的视频编码标准, 如H.26X系列。这些标准被广泛的应用于基于网络传输的视频通信, 极大的方便了人们对视频信息的获取。HEVC/H.265是继先进视频编码 (Advanced Video Coding, AVC/H.264) 之后最新一代视频编码标准, 其采用了许多新型技术来提高编码质量, 进一步满足了人们对高清, 甚高清视频的需求。但同时, 视频编码的计算复杂度限制了HEVC进一步发展。

[0003] 目前已经有许多算法被提出来通过牺牲一定的比特率来减少编码时间。主要可以分为减少帧内预测和帧间预测编码时间两个部分。如CN103327325中, 提出了基于HEVC标准的帧内预测模式快速自适应选择方法, 主要是通过对当前预测单元进行SATD (绝对误差和失真) 粗略选择, 比较候选模式中相邻两个代价差值与代价中值的比值与固定门限, 从而自适应选择候选模式, 再通过率失真优化来确定最佳预测模式。CN104954787B中, 提出了一种HEVC帧间预测模式选择方法, 该方法为: 确定当前帧的编码单元 (Coding Unit, CU) 块的像素尺寸在设定范围内时, 基于绝对误差和 (Sum of Absolute Differences, SAD) 算法计算获得所述CU块的每种帧间预测模式的第一率失真代价 (Sad Cost0), 根据所获得的各帧间预测模式的第一率失真代价获得最优的第一率失真代价及相应的第一帧间预测模式和次优的第一率失真代价 (Sad Cost1) 及相应的第二帧间预测模式, 确定 (Sad Cost1-Sad Cost0) / Sad Cost0大于阈值时, 选择所述第一帧间预测模式作为所述CU块的最优帧间预测模式。

[0004] 这些算法都减少了HEVC的编码复杂度, 但是我们知道, 在HEVC中, 耗时最长的部分是运动搜索部分。所以有效的减少运动搜索时间, 才能实现更加高效的视频编码。在HM标准代码中, 已经提出了TZ Search的方法来取代全搜索算法, 极大的减少了编码时间。运动搜索分为两步: 确定起始搜索点和确定最优搜索点。HEVC中采用高级运动矢量预测 (Advanced Motion Vector Predictor, AMVP) 技术来确定起始搜索点, AMVP会给出若干个候选预测运动矢量 (Motion Vector, MV), 编码器从中选择率失真代价最小的作为预测MV, 并用其所指向的位置作为起始搜索点。然后以步长1开始, 按照菱形模板 (如图1所示) 或正方形模板 (如图2所示) 在搜索范围内搜索, 其中以步长为2的整数次数幂的形式递增, 选出率失真最小的点作为该步骤的搜索结果, 然后补充搜索最优搜索点周围尚未搜索的点, 来确定最优搜索点。若该最优搜索点对应的步长大于某个阈值, 则以该最优点为中心, 在一定范围内做全搜

索,选择率失真代价最小的作为该步骤的最优点,以这个最优搜索点作为新的起始搜索点,细化搜索。当相邻向此细化搜索得到的最优点一致时停止细化搜索。此时得到的MV即为最终的MV。

发明内容

[0005] 本发明为克服上述现有技术所述的至少一种缺陷,提供一种基于HEVC图像内容复杂度的快速运动估计方法。本发明能够有效减少运动搜索损耗的时间,进而有效的改进HM中原有TZ Search的编码效率。

[0006] 为解决上述技术问题,本发明采用的技术方案是:一种基于HEVC图像内容复杂度的快速运动估计方法,其中,包括如下步骤:

[0007] S1.用Merge模式编码当前CU块;

[0008] S2.判断当前CU块的运动复杂度,若 $\max\{|d_1-d_2|, |d_2-d_3|, |d_3-d_1|\} < d_{th}$,且 $\max\{|\theta_1-\theta_2|, |\theta_2-\theta_3|, |\theta_3-\theta_1|\} < \theta_{th}$,则进行步骤S3,反之,则进行步骤S4;

[0009] 其中, d_1, d_2, d_3 分别为当前CU块到与其相邻的左、上、右上三个方向上已编码CU块的MV大小; $\theta_1, \theta_2, \theta_3$ 分别为 d_1, d_2, d_3 与水平线的夹角; d_{th} 和 θ_{th} 为设定的阈值;

[0010] S3.用Inter_2N×2N模式编码当前CU块,且只进行步长为2的菱形搜索;

[0011] S4.用Inter_2N×2N模式编码当前CU块,进行完整的TZ Search,并判断最优搜索点是否在初始搜索点附近,若在,进行步骤S5,若不在,则进行步骤S6;

[0012] S5.分别用Inter_2N×N, Inter_N×2N, Inter_N×N, Inter_2N×nU, Inter_2N×nD, Inter_nL×2N以及Inter_nR×2N模式编码当前CU块,进行步长为2的菱形搜索,并确定Inter_2N×N, Inter_N×2N, Inter_N×N, Inter_2N×nU, Inter_2N×nD, Inter_nL×2N以及Inter_nR×2N模式下的最优MV。

[0013] S6.分别用Inter_2N×N, Inter_N×2N, Inter_N×N, Inter_2N×nU, Inter_2N×nD, Inter_nL×2N以及Inter_nR×2N模式编码当前CU块,进行完整的TZ Search,并确定Inter_2N×N, Inter_N×2N, Inter_N×N, Inter_2N×nU, Inter_2N×nD, Inter_nL×2N以及Inter_nR×2N模式下的最优MV。

[0014] 通过统计可以知道,对于运动较为简单的图像,它的最优搜索点总是在起始搜索点附近或者为(0,0),因此我们就可以设定一个阈值来判断运动是否简单。设当前CU块相邻已编码CU(左,上,右上)的MV大小分别为 d_1, d_2, d_3 ,若 $\max\{|d_1-d_2|, |d_2-d_3|, |d_3-d_1|\}$ 的值小于某个阈值 d_{th} ,且MV的角度差值 $\max\{|\theta_1-\theta_2|, |\theta_2-\theta_3|, |\theta_3-\theta_1|\}$ 小于某个阈值 θ_{th} 时,我们认为图像运动较为简单,则当前inter_2N×2N模式的搜索省略为步长为2的菱形搜索。同时我们假设inter_2N×2N模式为父PU,其他模式为子PU。判断inter_2N×2N模式的最优搜索点是否在初始搜索点附近,若在,其子PU也仅仅只进行到步长为2的菱形搜索就终止,此时选出的最优搜索点即为该子PU最终MV所指向的点。若不在,则进行完整的TZ Search来确定最终的MV。

[0015] 我们假设运动估计的编码复杂度为 $ME_{complexity}$,其定义为:

[0016] $ME_{complexity} = (T_{ME}/T_{encoder}) \times 100\%$

[0017] 其中 T_{ME} 为HM源代码中运动估计的时间, $T_{encoder}$ 为HM源代码中整个编码过程所消耗的时间。

[0018] 表1为不同分辨率的视频序列在不同QP (Quantization parameter, 量化参数) 的情况下各运动估计的复杂度。

[0019] 表1运动估计的编码复杂度 $ME_{complexity}$ (%)

分辨率	序列	QP=22	QP=27	QP=32	QP=37	平均
416× 240	BQ Square	28.3	38.6	47.33	52.8	41.7
832× 480	Basketball Drill	36.3	42.8	48.8	53.5	45.4
1920× 1080	Cactus	45.3	49.3	52.6	65.9	53.3
2560× 1600	Traffic	41.4	48.9	53.8	56.7	49.3
	平均	57.12	65.94	71.17	74.65	66.47

[0021] 从表1中我们可以知道,运动估计在整个编码中消耗的时间比重非常大,因此有效的减少运动估计的时间,即可以大量减少编码时间。

[0022] 表2P(A)的数据结果 (%)

分辨率	序列	QP=22	QP=27	QP=32	QP=37	平均
416× 240	BQ Square	72.23	73.01	75.30	76.15	74.20
832× 480	Basketball Drill	81.32	83.06	84.69	85.71	83.70
1920× 1080	Cactus	91.72	92.31	94.56	95.02	93.40
2560× 1600	Traffic	71.01	72.23	73.51	75.69	73.11
	平均	79.07	80.15	82.02	83.14	81.10

[0024] 我们假设事件A为Inter_2N×2N模式中,其最优搜索点在初始搜索点附近或者为(0,0)的概率,实验得到P(A)的数据如表2所示。通过数据分析我们可以知道在运动较为简单的序列中,其概率可以高达95.02%,因此可以充分说明,对于提前终止TZ Search的方法是可行的。

[0025] 我们假设事件B为子PU的最优搜索点在初始搜索点附近或者为(0,0)。我们计算当

事件A成立时事件B发生的概率 $P(B|A)$, $P(B|A) = \frac{P(AB)}{P(A)} = \frac{C/N}{A/N}$, 其中C表示在运行过程中, 事件B发生的次数, A表示事件A发生的次数。通过统计可以发现, 其概率可以高达96.8%以上。因此可以充分说明, 如果父PU即Inter_2N×2N模式的最优搜索点在初始搜索点附近, 则其子PU (Inter_2N×N, Inter_N×2N, Inter_N×N, Inter_2N×nU, Inter_2N×nD, Inter_nL×2N, Inter_nR×2N) 的最优点搜索步骤可以仅仅只进行步长为2的TZ Search, 剩余步骤可以跳过。

[0026] 与现有技术相比, 本发明的有益效果:

[0027] 本发明的估计方法能够基于运动复杂度的判断来提前终止TZ Search, 以便进一步减少运动搜索损耗的时间, 进而可以有效的改进HM中原有TZ Search的编码效率。

附图说明

[0028] 图1是本发明中菱形搜索示意图。

[0029] 图2是本发明中正方形搜索示意图。

[0030] 图3是本发明中帧间预测PU划分示意图。

具体实施方式

[0031] 附图仅用于示例性说明, 不能理解为对本专利的限制; 为了更好说明本实施例, 附图某些部件会有省略、放大或缩小, 并不代表实际产品的尺寸; 对于本领域技术人员来说, 附图中某些公知结构及其说明可能省略是可以理解的。附图中描述位置关系仅用于示例性说明, 不能理解为对本专利的限制。

[0032] 一种基于HEVC图像内容复杂度的快速运动估计方法, 其中, 包括如下步骤:

[0033] S1. 用Merge模式编码当前CU块;

[0034] S2. 判断当前CU块的运动复杂度, 若 $\max\{|d_1-d_2|, |d_2-d_3|, |d_3-d_1|\} < d_{th}$, 且 $\max\{|\theta_1-\theta_2|, |\theta_2-\theta_3|, |\theta_3-\theta_1|\} < \theta_{th}$, 则进行步骤S3, 反之, 则进行步骤S4;

[0035] 其中, d_1, d_2, d_3 分别为当前CU块到与其相邻的左、上、右上三个方向上已编码CU块的MV大小; $\theta_1, \theta_2, \theta_3$ 分别为 d_1, d_2, d_3 与水平线的夹角; d_{th} 和 θ_{th} 为设定的阈值;

[0036] S3. 用Inter_2N×2N模式编码当前CU块, 且只进行步长为2的菱形搜索;

[0037] S4. 用Inter_2N×2N模式编码当前CU块, 进行完整的TZ Search, 并判断最优搜索点是否在初始搜索点附近, 若在, 进行步骤S5, 若不在, 则进行步骤S6;

[0038] S5. 分别用Inter_2N×N, Inter_N×2N, Inter_N×N, Inter_2N×nU, Inter_2N×nD, Inter_nL×2N以及Inter_nR×2N模式编码当前CU块, 进行步长为2的菱形搜索, 并确定Inter_2N×N, Inter_N×2N, Inter_N×N, Inter_2N×nU, Inter_2N×nD, Inter_nL×2N以及Inter_nR×2N模式下的最优MV。

[0039] S6. 分别用Inter_2N×N, Inter_N×2N, Inter_N×N, Inter_2N×nU, Inter_2N×nD, Inter_nL×2N以及Inter_nR×2N模式编码当前CU块, 进行完整的TZ Search, 并确定Inter_2N×N, Inter_N×2N, Inter_N×N, Inter_2N×nU, Inter_2N×nD, Inter_nL×2N以及Inter_nR×2N模式下的最优MV。

[0040] 通过统计可以知道, 对于运动较为简单的图像, 它的最优搜索点总是在起始搜索

点附近或者为(0,0),因此我们就可以设定一个阈值来判断运动是否简单。设当前CU块相邻已编码CU(左,上,右上)的MV大小分别为 d_1, d_2, d_3 ,若 $\max\{|d_1-d_2|, |d_2-d_3|, |d_3-d_1|\}$ 的值小于某个阈值 d_{th} ,且MV的角度差值 $\max\{|\theta_1-\theta_2|, |\theta_2-\theta_3|, |\theta_3-\theta_1|\}$ 小于某个阈值 θ_{th} 时,我们认为图像运动较为简单,则当前inter_2N×2N模式的搜索省略为步长为2的菱形搜索。同时我们假设inter_2N×2N模式为父PU,其他模式为子PU,如图3所示。判断inter_2N×2N模式的最优搜索点是否在初始搜索点附近,若在,其子PU也仅仅只进行到步长为2的菱形搜索就终止,此时选出的最优搜索点即为该子PU最终MV所指向的点。若不在,则进行完整的TZ Search来确定最终的MV。

[0041] 本实施例中,分别计算QP=22,27,32,37四个不同值的情况下,比较本发明算法与HM16.16标准代码TZ Search的编码时间,联合比特率(Bjontegaard Delta Bit Rate, BDBR)以及联合峰值信噪比(Bjontegaard Delta Peak Signal-to-Noise Ratio, BDPSNR),仿真平台为Visual Studio 2015,编码帧数为50, $d_{th}=1, \theta_{th}=15^\circ$ 。仿真结果如表3所示,通过对仿真结果分析可知,相比于HM中的TZ Search,本发明算法节省的运动估计时间约为13.44%,增加的BDBR约为0.26db,减少的BDPSNR约为0.014%。该数据充分说明了本发明算法可以有效的改进HM中原有TZ Search的编码效率。

[0042] 表3本发明算法结果/HM中TZ Search的结果

分辨率	序列	BDPSNR (%)	BDBR (db)	MTETS (%)
416×240	BQ Square	-0.011	0.56	-12.15
832×480	Basketball Drill	-0.029	0.22	-13.44
1920×1080	Cactus	-0.008	0.12	-10.69
1280×720	Four People	-0.007	0.14	-6.21
	平均	-0.014	0.26	-10.62

[0044] 显然,本发明的上述实施例仅仅是为了清楚地说明本发明所作的举例,而并非是对本发明的实施方式的限定。对于所属领域的普通技术人员来说,在上述说明的基础上还可以做出其它不同形式的变化或变动。这里无需也无法对所有的实施方式予以穷举。凡在本发明的精神和原则之内所作的任何修改、等同替换和改进等,均应包含在本发明权利要求的保护范围之内。

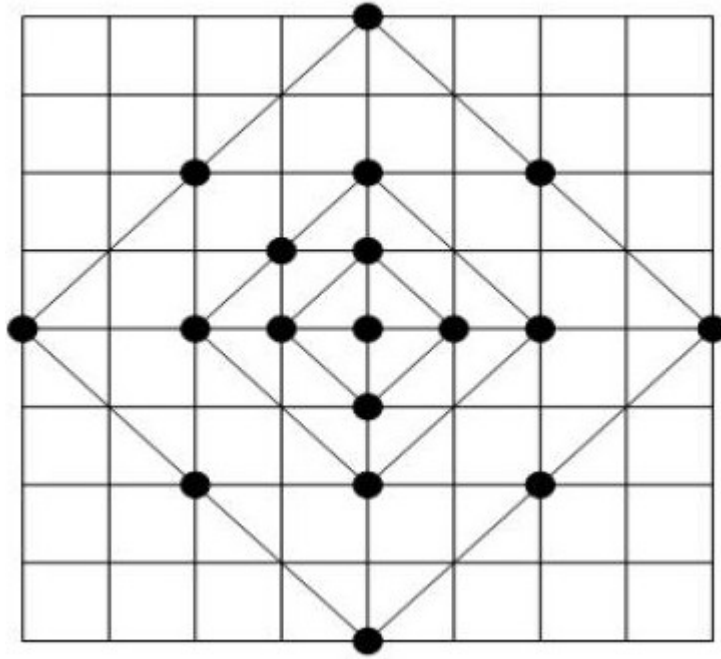


图1

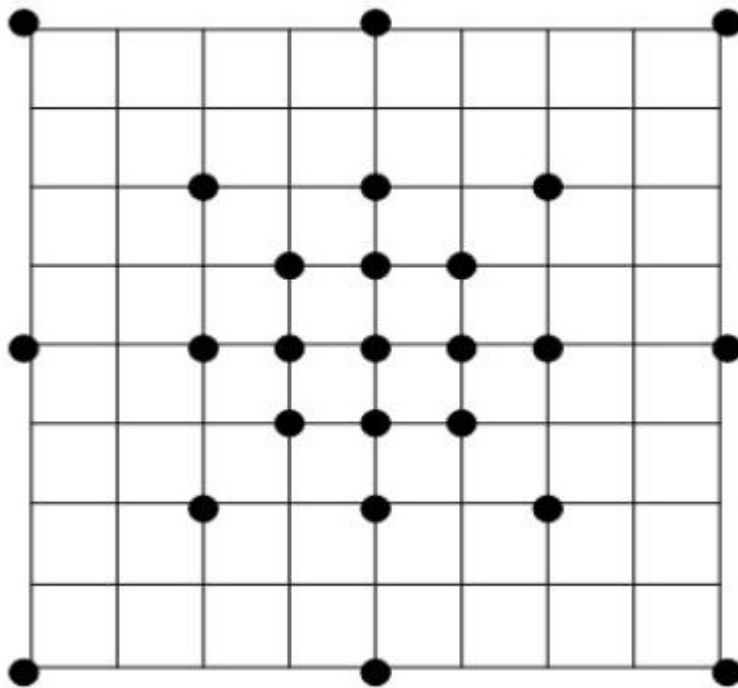


图2

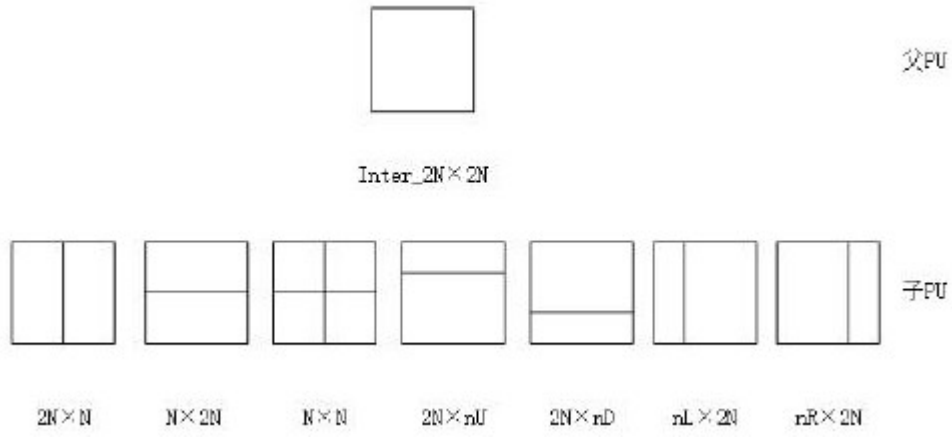


图3