

# 基于内容的匹配算法

李峰 周源华

(上海交通大学图像通信与信息处理研究所, 上海 200030)

TP391.41

**摘要** 从分析梯度法和相位法所采用的展开模型入手, 提出了基于内容的信号展开模型, 它是将信号的频谱进行搬移, 使其特征频率移至 0 频率处, 然后再利用 Taylor 公式将其展开, 从而避免了梯度法和相位法的缺点, 使此模型在用于信号展开时, 近似误差在更大的频带范围内保持恒定, 实现了采用此模型的匹配算法, 并且通过实验对比了梯度法、相位法和新算法的性能。

**关键词** 梯度法, 相位法, 基于内容, 展开模型

匹配算法 计算机视觉

## CONTENT-BASED MATCHING METHOD

LI Feng ZHOU Yuan-Hua

(Inst. of Image Communication & Information Processing, Shanghai Jiaotong University,  
Shanghai 200030, China)

**Abstract** Through analyzing expansion models of gradient-based method and phase-based method, a content-based signal expansion model was presented, in which the signal's spectrum is shifted to make its character frequency located at zero frequency, then the signal is expanded by Taylor formula, so the drawbacks of gradient-based method and phase-based method are avoided. When used in signal expansion, this model can give a relative constant error of approximation in a larger frequency band. The matching algorithm was realized based on the new expansion model, and the performance of gradient-based method, phase-based method and the new method was compared by experiment.

**Key words** gradient-based method, phase-based method, content-based, expansion model.

### 引言

图像匹配是计算机视觉中的一个基本问题, 例如, 立体视觉恢复、运动分析、特征检测和目标识别常常需要对图像进行匹配。此外, 一些非视觉问题, 如视频压缩、图像搜寻等也可以借助匹配来完成<sup>[1,2]</sup>。

匹配的方法是多种多样的, 梯度法和相位法是其中有代表性的两类方法, 梯度法的一般原理是将信号作线性展开(一般采用 Taylor 公式), 将信号中的视差变量分离出来, 然后根据某种准则(如最小均方误差准则等)构造目标函数并求解, 即可得到视差值<sup>[3]</sup>。此方法中展开式的系数通常是像素灰度的梯度或高阶导数, 故名梯度法。由于 Taylor 公式的限制, 此方法只能对低频信号很好地近似, 对于高频则会产生较大的误差, 因此图像中如果有很多变化明显而致密的细节时, 就会影响匹配的效果。相位法的一般原理是利用“信号在空域上的偏移等于其频域的相位产生相应的变化”这一特点, 对信号进行变换, 使视差分量成为频域中的

相位项, 据此可以求出视差<sup>[4]</sup>。通过分析发现: 与梯度相反, 相位法对低频信号会产生较大的误差, 而对于高频信号则有相对较好的效果。

本文提出了一种基于内容的匹配模型, 它综合了梯度法和相位法的优点而避免了它们的缺点。实验证明了基于此模型的匹配算法能够得到比梯度法和相位法更好的效果。

### 1 梯度法和相位法

为便于说明和实现的方便, 本文只考虑一维视差的情形。

设信号  $i_1(n), i_2(n)$  之间存在着视差  $d$ , 则有

$$i_1(n) = i_2(n + d). \quad (1)$$

对于梯度法而言, 首先利用 Taylor 公式将式(1)右边展开为

$$i_2(n + d) \approx i_2(n) + i_2'(n)d. \quad (2)$$

再根据最小均方误差准则, 得到目标函数

$$\sum_{n=-\infty}^{+\infty} w(n)(i_1(n) - (i_2(n) + i'_2(n)d))^2. \quad (3)$$

式(3)中  $w(n)$  是窗口函数, 式(3)是  $d$  的函数, 使式(3)取得极小值时的  $d$  即是所求的视差.

根据 Taylor 公式的性质, 式(2)只有当  $d$  很小时才有效. 显然, 对于相同的  $d$ , 用式(2)近似低频信号比近似高频信号的误差更小. 对式(2)两边做傅立叶变换, 有

$$e^{n\omega d} \approx 1 + \omega d. \quad (4)$$

显然, 式(4)只有当  $\omega \approx 0$  时, 等号两边才严格相等, 随着  $\omega$  的增加, 近似的误差急剧增加. 图 1 给出了当  $d=1$  时, 误差  $|e^{n\omega d} - (1 + \omega d)|$  随  $\omega$  变化的情况. 图中横轴  $\omega$  是频率, 纵轴  $V$  是误差.

通过以上分析, 可以看出梯度法对低频信号的匹配效果较好, 而对高频信号的匹配不理想, 所以梯度法对图像中致密的细节匹配效果不是很好.

相位匹配一般是先将  $i_2(n)$  乘以一个局域化函数(滤波器  $w(n)$ ), 然后进行傅里叶变换, 得到

$$\begin{aligned} Fw(n)i_2(n+d) &= (1/2\pi) \int_{-\pi}^{\pi} I_2(\theta)e^{j\theta d}W(\omega-\theta)d\theta \\ &\approx (1/2\pi)e^{j\omega d} \int_{-\pi}^{\pi} I_2(\theta)W(\omega-\theta)d\theta \\ &\approx e^{j\omega d}I_2(\omega). \end{aligned} \quad (5)$$

其中  $W(\omega) = F(w(n))$ , 相位法的目标函数为

$$\sum_{n=0}^N |I_1(\omega) - e^{j\omega d}I_2(\omega)|^2. \quad (6)$$

显然, 只有当  $W(\omega)$  是一个近似于  $\delta(\omega)$  的窄带滤波器时, 式(6)才有效. 但根据测不准原理, 滤波器的带宽近似与其时宽成反比, 在匹配中, 为了求出局域化的视差, 滤波器  $W(n)$  都是空间局域化的, 并且考虑到窗口效应, 滤波器的时宽不能太大, 因此其带宽一般都不会

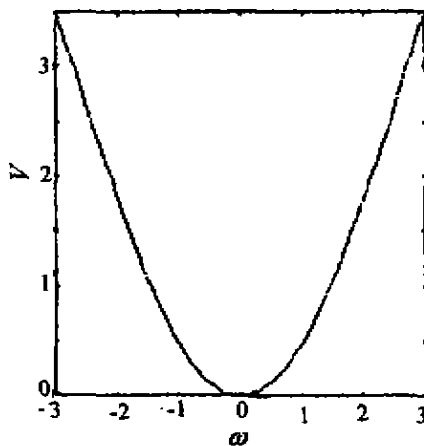


图 1 Taylor 公式的误差  
Fig. 1 Error of Taylor formula

太窄.

设峰值频率是  $\omega_p$ , 信号是频率为  $\omega$  单频信号, 则在式(5)的推导过程中隐含如下近似:

$$d_{\text{估}} = (\omega d) / \omega_p. \quad (7)$$

设带通滤波器  $W(\omega)$  的带宽是  $\Delta\omega$ , 式(5)中  $\omega$  的实际值可在  $[\omega - \Delta\omega, \omega + \Delta\omega]$ , 则视差估计的误差为

$$\Delta d_{\text{估}} = \pm (\Delta\omega d) / (2\omega_p). \quad (8)$$

图 2 给出了  $d=1.0$  和  $\Delta\omega=0.1$  时误差  $\Delta d_{\text{估}}$  随滤波器峰值频率  $\omega_p$  变化的情况, 可以看出相位法在低频段的近似将产生很大的误差. 横轴  $\omega$  是频率, 纵轴  $\Delta d$  是视差估计的误差.

## 2 基于内容的信号展开模型

设信号  $i_2(n+d)$  的中心频率是  $\omega_p$ , 则有

$$i_2(n+d) = (i_2(n+d)e^{-j\omega_p(n+d)})e^{j\omega_p(n+d)}. \quad (9)$$

式(9)中等号右边括号部分的中心频率是 0, 利用 Taylor 公式将式(9)展开为

$$\begin{aligned} i_2(n+d) &\approx (i_2(n)e^{-j\omega_p n}) + (i_2(n)e^{-j\omega_p n} - \\ &\quad j\omega_p i_2(n)e^{-j\omega_p n}d)e^{j\omega_p(n+d)} \approx e^{j\omega_p d}(i_2(n) + \\ &\quad (i_2'(n) - j\omega_p i_2(n))d). \end{aligned} \quad (10)$$

式(10)就是基于内容的信号展开模型, 其中  $i_2'(n)$  是  $i_2(n)$  的微分. 当  $\omega_p$  逼近于 0 时, 式(10)逼近于梯度法的信号展开模型式(2), 而当带通滤波器的带通很窄时, 式(10)逼近于相位法的信号展开模型式(5). 例如, 如果  $i_2(n)$  是一个单频函数, 式(10)则成为  $i_2(n+d) \approx e^{j\omega_p d}i_2(n)$ .

可见, 此模型的近似误差与频率是无关的, 无论信号的中心频率位于何处, 都可以得到较好的展开效果. 图 3 对 3 种展开模型进行了比较, 图中曲线 G、P 和 C

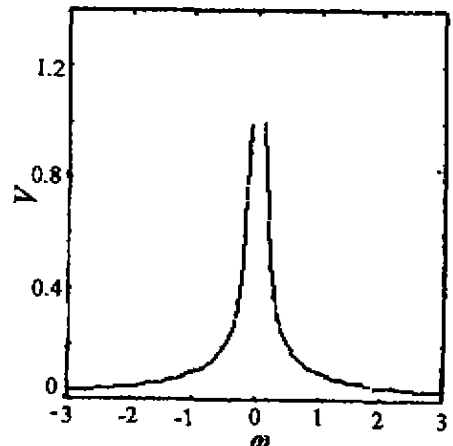


图 2 相位法的近似误差  
Fig. 2 Relative approximation error of phase-based method

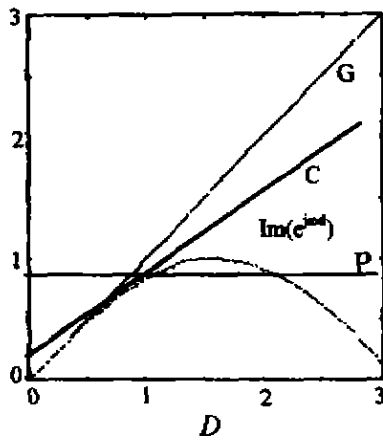


图3 3种信号展开模型的比较  
Fig. 3 Comparison of three signal expansion models

分别表示梯度法、相位法和基于内容的方法的展开方式对  $\text{Im}(e^{j\omega_p d})$  的逼近情况, 峰值频率  $\omega_p$  是 1.0. 很明显, 梯度法所采用的分解模型只在低频段有效, 在  $\omega_p$  附近的误差则比较大. 而相位法只在  $\omega_p$  点的效果较好, 在其周围的误差明显增加. 基于内容的模型在  $\omega_p$  及其周围的比较大的一段范围内都能够比较好地逼近信号.

由式(10)可以得到新的展开模型的目标函数为

$$\sum_N (i_1(n) - e^{-j\omega_p d} (i_2(n) + (i_2'(n) - j\omega_p i_2(n))d))^2, \quad (11)$$

式(11)中,  $N$  是窗口  $i_2(n)$  的宽度. 利用 Parseval 公式, 可以得到频域形式的目标函数

$$\sum_N (I_1(e^{j\omega k}) - e^{-j\omega_p d} (I_2(e^{j\omega k}) + (D_2(e^{j\omega k}) - j\omega_p I_2(e^{j\omega k}))d))^2. \quad (12)$$

其中  $\omega(k) = 2k\pi/N, k = 0, 2, \dots, N-1$ .

某些信号有一个“主要的”频率分量, 此信号的变化特性主要由此频率分量所决定, 则称此频率分量的频率为信号的特征频率. 设信号的特征频率为  $\omega_c$ , 中心频率(即滤波器的峰值频率)为  $\omega_p$ , 显然, 在对信号进行频率搬移时, 使  $\omega_c$  移至 0 频率处, 比使  $\omega_p$  移至 0 的频率处更为合理. 有些带通信号并不具有特征频率, 此时可以用峰值频率代替.

求解特征频率需要首先计算稳定性测度<sup>[5,6]</sup>

$$S = \frac{\sum_{n=0}^{N-1} i'(n)W(-n)}{\sum_{n=0}^{N-1} i(n)W(-n)}$$

其中,  $N$  是滤波器的长度, 然后计算判据  $|S - j\omega_p|$ ,  $\omega_p$  是滤波器的峰值频率. 此判据代表了滤波器的输出在多大程度上近似于一个正弦函数, 如果  $|S - j\omega_p|$  大于

给定的阈值, 则认为没有特征频率, 否则, 取  $\omega_c = \text{Im}(S)$  为特征频率. 求得特征频率后, 用  $\omega_c$  代替式(10)、(11)中的  $\omega_0$  即可.

### 3 基于内容的图像匹配算法

基于前面的分析, 我们给出采用基于内容的展开模型的匹配算法: (1) 通过窗口函数  $W(n)$ , ( $n = 0, 1, \dots, N-1$ ) 在左、右图像上截取  $i_1(n)$  和  $i_2(n)$ , 计算  $i_2(n)$  的灰度梯度  $i_2'(n)$ ; (2) 对  $i_1(n)$ ,  $i_2(n)$  和  $i_2'(n)$  进行 Fourier 变换, 得到  $I_1(e^{j\omega k})$ ,  $I_2(e^{j\omega k})$  和  $D_2(e^{j\omega k})$ ; (3) 计算  $i_2(n)$  的特征频率  $\omega_c = \text{Im}(D_2(e^{j\omega k})/I_2(e^{j\omega k}))$  和其稳定性测度  $T = |\text{Im}(D_2(e^{j\omega k})/I_2(e^{j\omega k})) - j\omega(k)|$ , 如果  $T > T_0$ , 则认为没有特征频率, 令  $\omega_c = \omega_p$ , 否则,  $\omega_c$  是特征频率; (4) 通过使下式最小来计算视差  $d$ :

$$\sum_N (I_1(e^{j\omega k}) - e^{-j\omega d} (I_2(e^{j\omega k}) + (D_2(e^{j\omega k}) - j\omega_p I_2(e^{j\omega k}))d))^2, \quad (13)$$

算法中的窗口一般不能采用矩形窗, 因为矩形窗的频谱的旁瓣较大, 且延伸很远, 会影像算法的效果. 本文采用 Blackman 窗, 其形式为

$$W(n) = (0.42 - 0.5\cos(2\pi n/(N-1)) + 0.08\cos(4\pi n/(N-1)))R_N(n). \quad (14)$$

通常, 灰度梯度的计算是采用  $[-101]/2$  这种简单的滤波器来完成的, 然而其性能并不好, 理想的梯度滤波器的频率, 从图可以看出仅仅在低频部分的一小段内,  $[-101]/2$  接近于理想的梯度滤波器. 为了提高梯度求取的精度, 本文采用如下的滤波器<sup>[7]</sup>:

$$[0.008852434 - 0.03666318 \ 0.09214035 - 0.1956004 \ 0.3975202 - 0.9452388 \ 0 \ 0.9452388 - 0.3975202 \ 0.1956004 - 0.09214035 \ 0.03666318 - 0.008852434],$$

此梯度滤波器的性能比  $[-101]/2$  等要好得多.

### 4 实验

以 PC 133+Win95 为平台, 用 Visual C++ 分别对梯度法、相位法和本文算法进行了模拟. 实验所用图像为 pentagon 像对, 是  $512 \times 512 \times 8$  bits. 实验结果如图 4~6 所示.

Pentagon 像对的视差范围为  $-10 \sim 10$  (pixel), 所以要结合从粗至精的多尺度匹配策略. 本文结合金字塔分解来进行, 上一层求出的视差用于指导下一层的匹配. 图像共分为 3 层(由粗至精), 并且只对列进行分解:  $512 \times 128 \rightarrow 512 \times 512$ , 窗口为 16 pixels, 假如设定

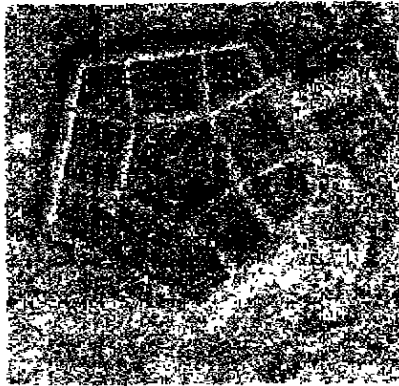


图 4 Pentagon 立体像对  
Fig. 4 Stereo pair of pentagon

最大视差为不超过窗口大小的  $1/4$ , 即 4 pixels. 则通过 3 层金字塔分解后能够求取的最大视差为  $4 \times 2^{3-1} = 16$  pixels. 阈值  $T$  取 0.5. 从图 4~6 可以看出采用基于内容的匹配算法对图像细节的匹配要好于梯度法和相位法. 表 1 给出了对经过校正的某航空像对 (2361  $\times$  2152 pixels) 中的 3 个地面控制点 (事先测定的匹配点) 进行匹配的精度. 点位置的单位是 pixel, 右点的估计位置是左点的位置加上视差得到的, 由右点的估计位置减去右点的真实位置就得到视差估算的误差.

表 1 3 个控制点的匹配精度

Table 1 The matching precision of three control points

控制点 (编号)	右点位置	右点位置 (真实)	右点位置 (估计)	误差 (pixel)
1274	1274	432	432.745	+0.745
2178	1271	425	426.064	+1.064
2179	2085	1231	1231.007	+0.007



图 5 梯度法所求视差的灰度图  
Fig. 5 Intensity of disparities obtained by gradient-based method



图 6 相位法所求视差的灰度图  
Fig. 6 Intensity of disparities obtained by phase-based method

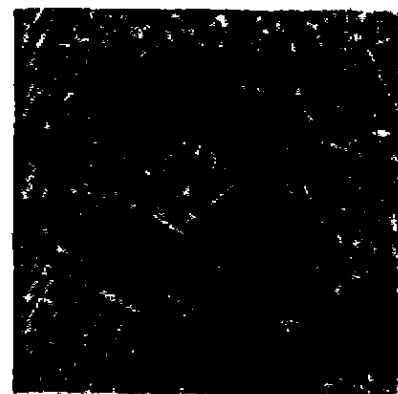


图 7 新算法所求视差的灰度图  
Fig. 7 Intensity of disparities obtained by the new method

## 5 结论

本文通过分析梯度法和相位法的展开模型,发现它们在性能上互补的特点,并进而提出了基于内容的信号展开模型,综合了梯度法与相位法的展开模型的优点而避免了它们的缺点,理论和实验都证明了采用基于内容的展开模型的匹配算法,能够得到比梯度法和相位法更好的效果.

## REFERENCES

- [1] Alparone V L. Digital processing of stereo images and 3-D reconstruction techniques, *International Journal of Remote Sensing*, 1991, **12**(3):477-490
- [2] ZHANG Zu-Xun, ZHANG Jian-Qing. *Digital Photogrammetry*, Wuhan: Press of Wuhan Technology University of Surveying and Mapping(张祖勋,张剑清. 数字摄影测量学, 武汉: 武汉测绘科技大学出版社),1996
- [3] Lucas B. Generalized image matching by the method of difference, PhD thesis. Department of computer science, Carnegie Mellon University. 1984
- [4] Fleet D J, Jepson A D. Phase-based disparity measurement. *CVGIP: Image Understanding*,1991, **53**(2): 198-210
- [5] Fleet D, Jepson A. Stability of phase information. *Proc. of IEEE Workshop on Visual Motion*, New Jersey, Princeton, October,1991: 52-60
- [6] Yalin Xiong, Shafer S A. Variable window Cabor filters and their use in focus and correspondence, *Proc. Computer Vision and Pattern Recognition*, 1994: 668-671
- [7] Sunder S. On the efficient design of digital differentiators. *International Journal of Circuit Theory and Applications*, 1994, **22**: 335-340