

9. ANÁLISE POR PRINCIPAIS COMPONENTES

9.1 Introdução

A **Análise por Principais Componentes (APC)**, também chamada de Transformação por Principais Componentes ou Transformada de Karhunen-Loeve, é uma das funções de processamento de imagens mais poderosas na manipulação de dados multiespectrais. Ao mesmo tempo, ela é também uma das menos compreendidas e, precisamente por isso, talvez uma das mais subutilizadas das técnicas de processamento digital de imagens. Contudo, com o aumento crescente no número de bandas espectrais disponíveis nos modernos sensores, a APC vem se tornando cada vez mais relevante na análise espectral de imagens multiespectrais.

Um dos fatores que contribuiu para o uso inadequado dessa técnica no processamento de imagens de S.R. está ligado ao próprio histórico da mesma. A aplicação da APC ao processamento de imagens iniciou-se no âmbito da Engenharia Elétrica, com a finalidade de **processar sinais eletrônicos para remoção de ruído**. Esse enfoque persistiu de certo modo quando a técnica passou a ser adotada no processamento de imagens de S.R., causando o aparecimento de "receitas" para o seu uso, que acabaram limitando o verdadeiro potencial da mesma, como será visto em detalhes mais adiante. Por outro lado, os conceitos matemáticos envolvidos na APC são consideravelmente complexos, o que contribuiu para que os princípios básicos por trás da técnica não fossem adequadamente compreendidos pelos usuários do processamento digital de imagens.

Outra característica da APC é que seus resultados são de difícil generalização para determinados tipos de aplicações. Na realidade, esses resultados são extremamente dependentes das características espectrais da cena sendo processada, não devendo ser transportados para outras situações.

A melhor maneira de se entender os conceitos envolvidos na APC, sem necessariamente aprofundar-se na definição e nos detalhes matemáticos da técnica, é usando o conceito de **espaço de atributos**, visto anteriormente quando se discutiu operações aritméticas e classificação.

Como já comentado anteriormente, as imagens produzidas por sensores remotos do tipos "*scanners*" possuem um alto grau de correlação entre as suas bandas multiespectrais, de tal forma que seus espaços de atributos são relativamente estreitos. A principal função da APC é **determinar a extensão dessa correlação e, através de uma transformação matemática apropriada, removê-la**. Um conjunto de N imagens de entrada vai produzir, quando processadas através da APC, um outro conjunto de N imagens de saída com nenhuma correlação entre si.

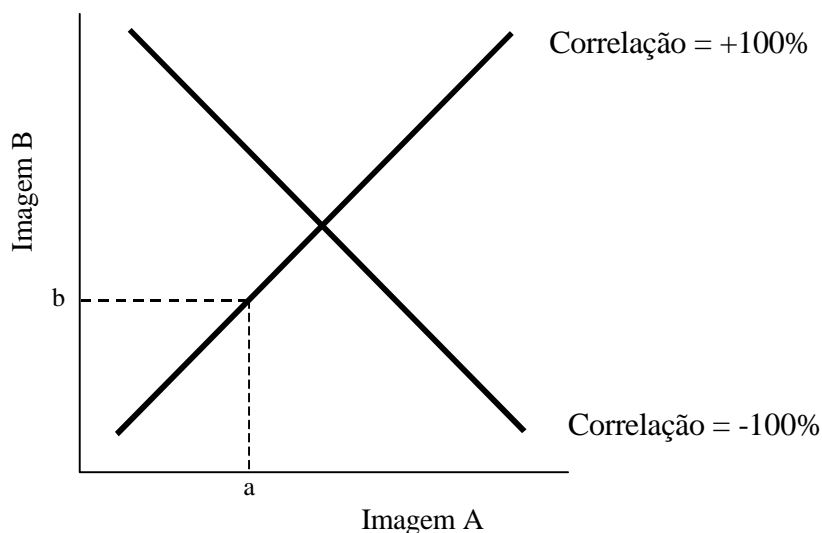
A Análise por Principais Componentes é apropriada no tratamento de imagens produzidas por sensores com um alto número de bandas espectrais. Ela é também de grande utilidade na análise de dados comumente utilizados em

geologia, tais como radar, geoquímica ou geofísica, em conjunto com imagens multiespectrais.

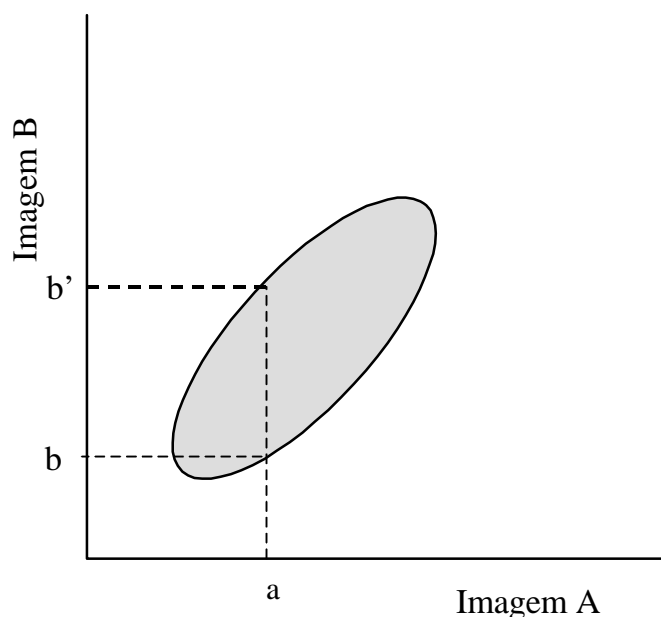
9.2 Correlação de imagens

Antes de iniciar a discussão da APC propriamente dita, convém analisar a razão e as implicações da correlação entre as bandas espectrais de uma imagem. **Duas imagens são ditas correlacionadas quando, dada a intensidade de um determinado pixel em uma delas, pode-se deduzir com razoável aproximação a intensidade do pixel correspondente na outra imagem.** Se as duas imagens são absolutamente idênticas, as imagens são ditas 100% correlacionadas. Se uma delas é o negativo da outra, essa dedução também pode ser feito precisamente, mas neste caso as imagens são ditas serem -100% correlacionadas.

Em um gráfico de espaço de atributos, a correlação entre imagens se mostra como uma distribuição de pontos ao longo da reta de direção próxima à diagonal principal. Em um caso bidimensional, duas imagens perfeitamente correlacionadas vão ter seu espaço de atributos representado por uma linha reta, inclinada a 45°, como mostrado na Figura 9.1a. O caso em que a correlação é -100% também é representado por uma reta, mas agora com uma inclinação negativa. Já um exemplo em que duas imagens apresentam uma correlação parcial é mostrado na Figura 9.1b, em que a distribuição tem uma forma elipsoidal ao longo da reta, mas agora com espalhamento finito.



a) Duas imagens (A, B) correlacionadas



b) Duas imagens com correlação menor do que 100%

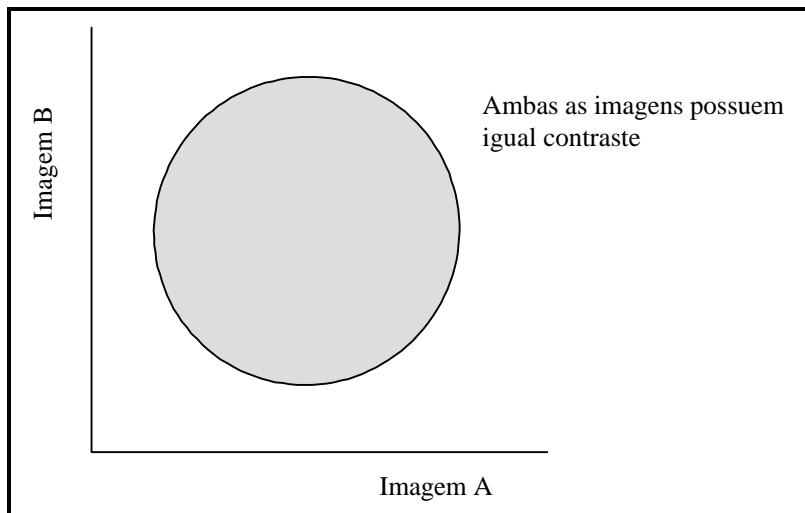
Figura 9.1 - Correlação em um espaço de atributos bidimensional

Comparando-se esses dois casos pode-se notar **como muda o processo de predição do valor da intensidade do pixel de uma imagem a partir do valor da outra**. No caso da Figura 9.1a, tomado um pixel da imagem A, acha-se um único pixel correspondente na imagem B. No caso da Figura 9.1b, acha-se um intervalo de valores de intensidade na imagem B correspondendo a um único pixel da imagem A; esse intervalo será inversamente proporcional à correlação entre as imagens.

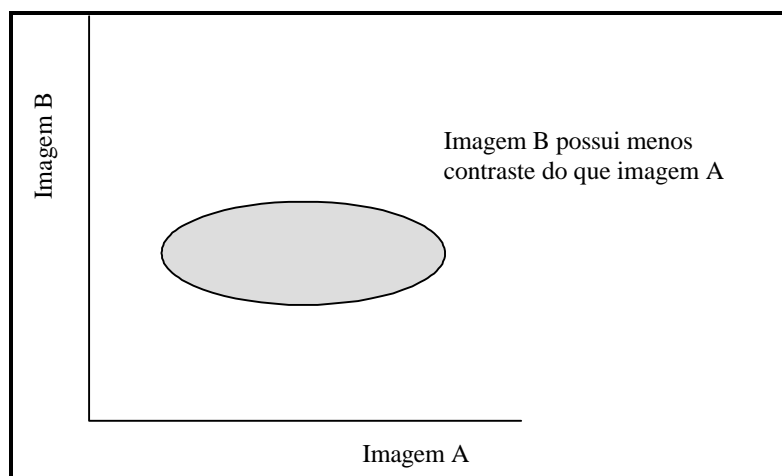
A Figura 9.2 mostra dois exemplos de espaços de atributos para imagens totalmente não-correlacionadas. O primeiro deles é uma extensão da Figura 9.1b, mostrando uma distribuição circular de pontos, onde o processo de predição não funciona (para um ponto da imagem A, é impossível prever qual o ponto correspondente em B). A Figura 9.2b mostra uma distribuição semelhante à da Figura 9.2a, mas com a imagem B tendo sofrido uma compressão em seu contraste, de tal forma que o intervalo total de intensidades foi consideravelmente reduzido. Da mesma forma, o processo de predição neste caso é também impossível.

Como já mencionado antes, as **causas da existência de correlação** podem ser várias, as mais importantes sendo o sombreamento topográfico e a relação entre o espectro de reflectância dos materiais superficiais e as bandas espectrais do sensor. Uma vez que praticamente todos os sensores multiespectrais dependem da iluminação solar, essa correlação vai quase sempre existir, com o sombreamento tendendo a afetar todas as bandas da mesma

maneira. O sombreamento, embora útil para a percepção dos detalhes texturais de uma imagem, pode obscurecer feições espectrais de interesse. Como exemplo, pode-se mencionar a dificuldade de classificar um determinado tipo de superfície localizada em uma área de sombra, quando a área de treinamento usada encontrava-se iluminada. Existem duas exceções nessa questão do sombreamento afetando igualmente todas as bandas: a primeira diz respeito às **imagens termais** (a banda 6 do TM, por exemplo) que, por representarem energia emitida, e não a refletida pelo sol, possuem muito menos correlação com as outras bandas refletidas. Outro caso são as **bandas de comprimento de onda na faixa da radiação visível de baixo comprimento de onda**, como a banda 1 do TM, que, por sofrerem intensamente os efeitos de espalhamento atmosférico, os quais não afetam as outras bandas da mesma forma, geralmente apresentam uma menor correlação com as outras bandas refletidas.



a)



b)

Figura 9.2 - Espaço de atributos de imagens não-correlacionadas

A correlação entre bandas multiespectrais, quaisquer que sejam suas causas, vai sempre obscurecer diferenças sutis entre reflectâncias dos materiais superficiais, dificultando a sua identificação e, conseqüentemente, a sua separação. Uma vez que correlação significa que uma imagem pode ser (mais ou menos) inferida a partir de outra, isto também implica que uma certa **redundância dos dados** ou, visto de um outro ângulo, um certo desperdício na manipulação, processamento e armazenamento de dados. Uma questão que se coloca neste ponto é: **"Se uma grande proporção de cada banda espectral é composta por informação redundante, por que não extrair essa informação comum, concentrando-a em uma única imagem ao invés de replicá-la várias vezes?"** O que restar após a extração da informação comum pode então ser distribuído em novas imagens. Esta noção forma a **base conceitual da APC**.

9.3 Variância e covariância de imagens

Para poder analisar as correlações entre as várias bandas de uma imagem, faz-se necessário alguma forma de representação matemática. Essa representação pode ser feita em função de alguns parâmetros, tratados a seguir.

Considerando o espaço de atributos bidimensional das duas imagens **A** e **B** vistas nas duas figuras anteriores, o parâmetro mais simples de especificar é a **posição do centróide da distribuição**. A Figura 9.3 mostra a posição do centróide, que representa o centro de gravidade da distribuição (e não o centro geométrico), significando o ponto em que um hipotético "pixel de intensidade média" da imagem seria localizado.

Uma vez que o centróide é conhecido, o tamanho da distribuição necessita ser especificado. Isso é feito através do cálculo do espalhamento ao longo de cada eixo, que recebe o nome de **variância**. A variância é definida pelo quadrado do desvio-padrão do histograma de intensidades para cada banda e dá a **medida do contraste de cada banda espectral**.

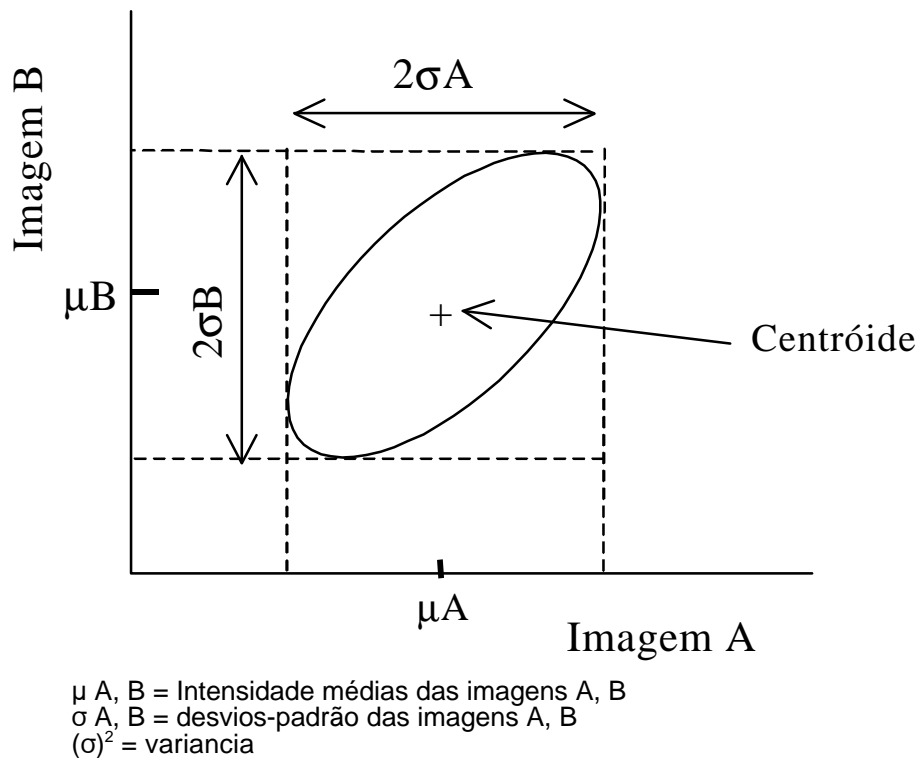


Figura 9.3 - Parâmetros estatísticos básicos do espaço de atributos.

Por último, a correlação entre cada par de bandas da imagem deve ser especificada, usando-se o parâmetro chamado de **covariância**. A covariância é uma grandeza que relaciona duas variâncias específicas e descreve o quanto a distribuição é espalhada ao longo da direção diagonal e, portanto, o grau de correlação existente entre duas bandas. Em outras palavras, ela dá a medida do quanto de informação contida em um par de bandas é comum a ambas. A covariância pode tanto ser positiva, e neste caso é dito que os dados são positivamente correlacionados, como negativa, quando eles são negativamente correlacionados.

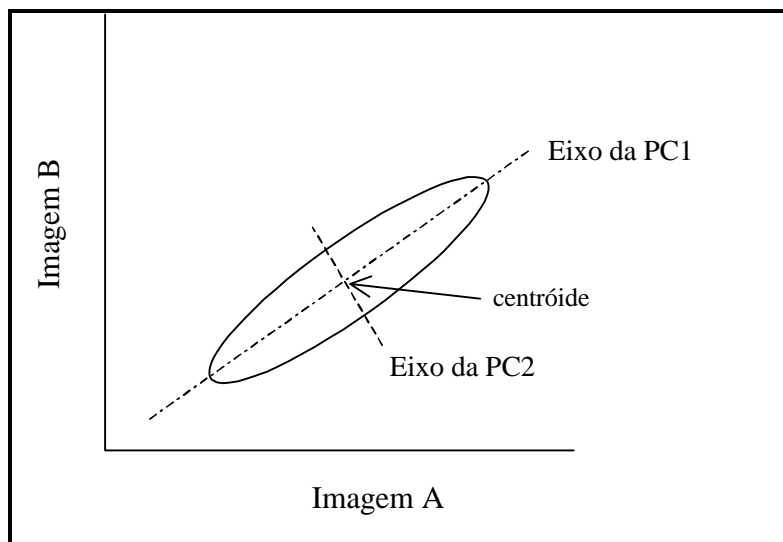
Na APC, é comum definir a covariância em termos do **coeficiente de correlação**. Esse coeficiente é calculado através da normalização da covariância, para considerar o espalhamento de contraste em cada banda individualmente. Isso é feito através da divisão da covariância pelo produto dos dois desvios-padrão (A,B).

As fórmulas matemáticas para o cálculo desses parâmetros requerem a inspeção de cada pixel de cada banda da imagem a ser processada, motivo pelo qual a APC é um processo bastante demorado em termos computacionais. Por exemplo, para uma imagem TM com 7 bandas, todos os aproximadamente 6 milhões de pixels são usados para se calcular 7 médias, 7 variâncias e 21 covariâncias!

9.4 A Transformação por Principais Componentes

A principal tarefa executada pela APC pode ser resumida da seguinte forma: dadas N imagens brutas (sem qualquer outro tipo de processamento), são analisadas as correlações entre elas e produzido um novo conjunto de N novas imagens, com nenhuma correlação entre si. De modo geral, esse novo conjunto de imagens terá as seguintes características: a primeira delas, chamada de 1ª Principal Componente, ou PC1, irá conter a informação que **é comum a todas as N bandas originais**; a 2ª PC irá conter a feição espectral mais significativa do conjunto e assim sucessivamente as PCs de ordem mais alta irão conter feições espectrais cada vez menos significantes, até a última PC, que concentrará então a informação que sobrar.

Para se entender como funciona esse agrupamento de dados baseados na correlação entre bandas, é necessário examinar uma representação gráfica da transformação por principais componentes. A Figura 9.4 mostra um espaço de atributos bidimensional (notar que isto é uma simplificação de uma situação real, onde o espaço de atributos terá tantas dimensões quantos forem as bandas espectrais consideradas!), com o centróide assinalado. Deve-se ser enfatizado que a **tendência alongada da distribuição é causada pela correlação entre as duas imagens** e que **o espalhamento em volta da diagonal é originada pelas diferenças espectrais entre as duas imagens** (quanto maior a diferença, mais longe da diagonal estarão os pontos da distribuição).



a) Determinação bidimensional dos eixos das PCs

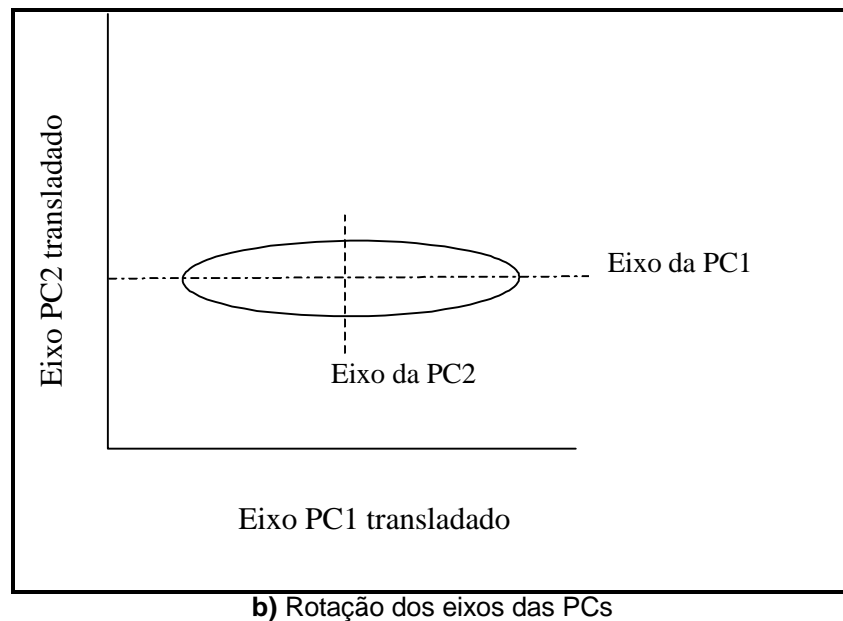


Figura 9.4 - Transformação por Principais Componentes

A transformação por principais componentes funciona da seguinte forma: primeiramente, **determina-se a direção relativa ao centróide que possui a máxima variância ou espalhamento, atribuindo-se à ela o eixo da 1ª PC.** Na Figura 9.4a essa direção é representada pelo eixo principal da elipse; essa direção descreve a correlação entre as imagens A e B. Em seguida, **determina-se a direção relativa ao centróide que possui a segunda maior variância ou espalhamento, com a condição de que ela deva ser ortogonal em relação à primeira direção, atribuindo-se à esta segunda direção o eixo da 2ª PC.** É claro que, neste exemplo bidimensional, não há nenhuma outra possibilidade para se colocar o eixo da 2ª PC, devido à restrição da ortogonalidade. Porém, em um caso real, usando por exemplo uma imagem TM com 6 bandas, existiriam pelo menos 5 direções alternativas para o eixo da 2ª PC após se ter determinado o eixo da 1ª PC. Esse processo é então repetido tantas vezes quanto for o número de bandas da imagem, até que todos os eixos tenham sido determinados. Por último, **rotaciona-se a distribuição do espaço de atributos em volta do centróide, até que cada uma das direções encontradas fique paralela a um dos eixos originais.** Essa distribuição rotacionada é mostrada na Figura 9.4b e representa as duas novas Principais Componentes calculadas a partir das imagens A e B.

A distribuição mostrada na Figura 9.4b é equivalente àquela da Figura 9.2b, podendo-se observar que não existe correlação alguma. No exemplo da Figura 9.4b, a correlação foi propositalmente removida pela rotação dos eixos, não havendo nenhuma tendenciosidade diagonal na distribuição.

Na prática, esse processo, acima descrito de maneira gráfica, é implementado algebricamente, usando-se os coeficientes de correlação ou as covariâncias para se determinar um conjunto de quantidades conhecidas por **auto-valores** (*eigenvalues*, em inglês). Os auto-valores representam o **comprimento dos eixos das PCs de uma imagem** e são medidos em unidades de variância.

Associados com cada auto-valor, existe um outro conjunto de quantidades, que é na verdade um conjunto de coordenadas, chamados de **auto-vetores** (*eigenvectors*), **que representam as direções dos eixos das PCs**. Os auto-vetores servem com fatores de ponderação que definem a contribuição de cada banda original para uma PC, numa combinação aditiva e linear que tem a seguinte forma:

$$\begin{aligned} \text{PC1} &= (\mathbf{B}_1 * \mathbf{e}_{1,1}) + (\mathbf{B}_2 * \mathbf{e}_{1,2}) + \dots + (\mathbf{B}_n * \mathbf{e}_{1,n}) \\ &\cdot \\ &\cdot \\ &\cdot \\ \text{PCN} &= (\mathbf{B}_1 * \mathbf{e}_{n,1}) + (\mathbf{B}_2 * \mathbf{e}_{n,2}) + \dots + (\mathbf{B}_n * \mathbf{e}_{n,n}) \end{aligned}$$

onde **PC1,...,PCN** são as principais componentes, **B1,...,B_n** são as bandas originais e **e_{1,1},...,e_{1,n}/e_{n,1},...,e_{n,n}** são os auto-vetores.

9.5 Aumento de Contraste e Composições Coloridas de Principais Componentes

Ao se observar a Figura 9.4b, nota-se que o espaço de atributos das imagens PC não é muito grande. Em particular, a segunda PC possui muito menos contraste do que a primeira, significando que ela possui muito menos informação. A quantidade de informação em cada PC é fornecida pelo respectivo auto-vetor. Isso é comum em APC de imagens multiespectrais, onde **cada sucessiva PC possui menos informação do que suas predecessoras**. Deve ser enfatizado que isso não quer dizer que essa informação seja qualitativamente menos importante, mas apenas que em termos quantitativos o conteúdo em informação decresce para as últimas PCs. Por outro lado, se as imagens PCs forem submetidas a um aumento de contraste, no sentido de se obter uma mesma variância para a primeira e a segunda PCs, o espaço de atributos resultante vai ser consideravelmente ampliado. Considerando o exemplo bi-dimensional da Figura 9.4b, obtém-se a distribuição mostrada na Figura 9.5.

Neste caso, foi aplicado um aumento de contraste independente nas duas PCs, o que aumentou seus espalhamentos ao longo dos dois eixos, tornando a distribuição final um círculo, e não mais uma elipse. Dessa forma, imagens PCs

com aumento de contraste utilizam muito melhor o espaço de atributos, permitindo uma melhor diferenciação dos materiais superficiais, uma vez que as nuvens de pontos que os representam irão plotar mais separadas nesse espaço.

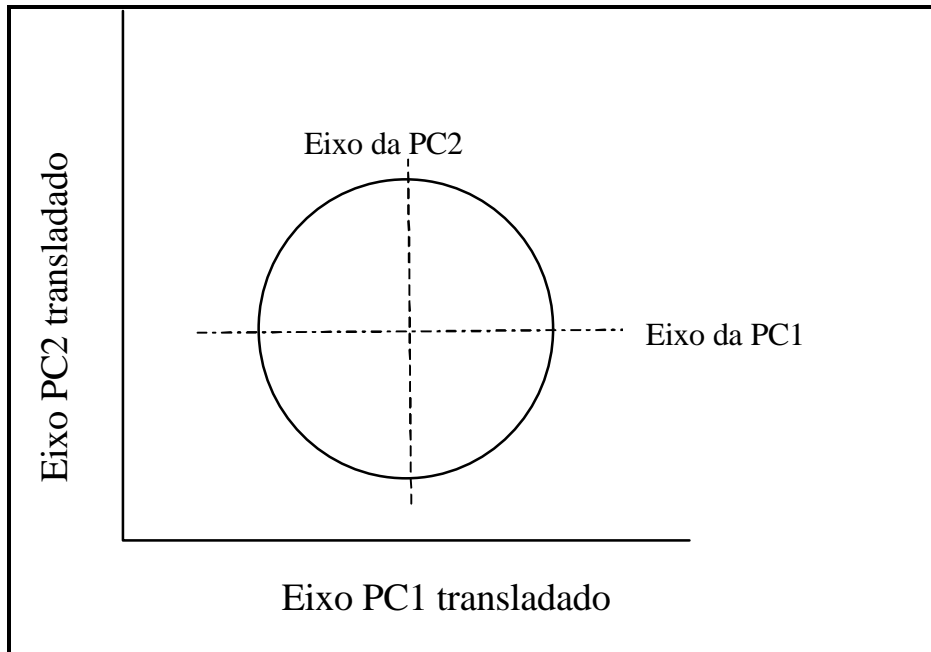


Figura 9.5 - Aumento de contraste aplicado em PCs para equalização da variância.

Deve ser notado que não é possível se obter o formato circular mostrado na figura 9.5 utilizando-se apenas técnicas de aumento de contraste em imagens originais, pois nesse caso tende-se apenas a "esticar" a distribuição elíptica ao longo dos eixos originais.

Imagens PCs podem ser combinadas em cores como quaisquer outras. A diferença é que a distribuição de cores em composições coloridas de PCs é a melhor possível, uma vez que não há correlação entre as imagens. Desse modo, uma composição colorida desse tipo praticamente não possui tons de cinza (o que indicaria correlação) mas apenas cores espectrais puras e intensamente saturadas.

9.6 Interpretação de imagens PCs

No tocante à interpretação de imagens Pcs, trata-se de uma tarefa razoavelmente complexa. O resultado da APC depende fortemente da forma exata da distribuição do espaço de atributos. Portanto, não há "receitas" fixas para a interpretação de Pcs e cada caso deve ser considerado separadamente. Uma das únicas regras que pode ser adotada é que a 1ª PC irá sempre conter o

albedo mais o sombreamento topográfico de uma cena de satélite, ou seja, é uma imagem muito semelhante à uma vista "pancromática" da cena. Todas as demais Pcs são completamente desprovidas de informação topográfica, devido à ausência do sombreamento.

Uma maneira de se compreender como as Pcs são formadas é examinando-se cuidadosamente a matriz de auto-vetores, para entender a contribuição de cada banda original na formação de cada uma das Pcs. Para tornar mais fácil a percepção dessas contribuições, deve-se transformar os auto-vetores em porcentagens, o que não modifica o resultado uma vez as Pcs são uma combinação aditiva linear das bandas originais. Uma vez calculadas as porcentagens e sabendo se o sinal de cada coeficiente do auto-vetor é positivo ou negativo, pode-se comparar com as curvas espectrais de materiais conhecidos (por exemplo, vegetação, água, diferentes tipos de solos e rochas, etc), determinando-se assim com um grau razoável de certeza em qual(is) PC(s) a informação espectral desejada irá ser concentrada.

Um outro ponto sobre Pcs é a questão da concentração de ruído. Já foi visto que o ruído, inerente à maioria dos sensores, representa informação não-correlacionada entre bandas. Assim, o ruído irá sempre ser extraído do conjunto de bandas multiespectrais pela APC e concentrado nas últimas PCs. Por esse motivo, uma das recomendações comuns em processamento de imagens multiespectrais é que se descarte as últimas PCs de um conjunto, para se livrar do ruído que irá se concentrar nelas. Assim, é comum encontrar, por exemplo, aplicações usando 3 das 4 primeiras PCs de uma imagem TM (que tem 6 bandas refletidas e, portanto, 6 PCs). O problema nesse caso é que, embora do ponto de vista quantitativo a porcentagem de informação nas duas últimas Pcs seja muito inferior a das anteriores, do ponto de vista qualitativo essa pequena porcentagem pode corresponder justamente à informação de interesse do usuário! Esse é o caso, por exemplo, de feições espectrais relacionadas a determinados minerais indicativos da presença de jazidas, que, por se tratar de feições espectrais muito específicas, fica normalmente concentrado nas 2 ou 3 últimas PCs!

9.7 Modificações da APC

O processo da APC pode ser modificado de algumas formas, no sentido de melhorar sua aplicação em alguns casos específicos. Existem duas modificações da APC que são comumente utilizadas. A primeira delas é denominada **Análise Canônica** e a segunda **Realce por Decorrelação**.

9.7.1 Análise Canônica

Em termos conceituais, a Análise Canônica representa uma mistura de classificação e principais componentes, funcionando da seguinte maneira: se numa cena sabe-se que existem determinados tipos de coberturas superficiais e se a posição das nuvens de pontos que os representam no espaço de atributos é mais ou menos conhecida, então **a APC pode ser direcionada de tal forma que tudo que não pertencer a esses tipos de coberturas será ignorado quando da aquisição dos parâmetros estatísticos**. Esse direcionamento é feito através da determinação das áreas nas quais se quer que a estatística seja considerada, usando a mesma sistemática de definição de "áreas de treinamento" empregada na classificação. No espaço de atributos esse procedimento vai remover as nuvens correspondentes aos tipos de cobertura sem interesse, deixando apenas as nuvens dos materiais superficiais de interesse. Em uma situação em que há uma superposição, no espaço de atributos, entre as nuvens dos tipos de cobertura de interesse com outras nuvens sem interesse, esse processo é bastante eficaz.

Um exemplo de situação em que o emprego da Análise Canônica pode ser útil é quando se tem uma área com manchas de cobertura vegetal e se quer processar a imagem para realce de feições geológicas ou de solos, que normalmente seriam obscurecidas pela presença da vegetação. Neste caso, pode especificar como "área de treinamento" para o cálculo dos parâmetros estatísticos (matriz de covariância) da APC as regiões onde as rochas (ou solos) são bem expostos, ignorando as áreas com cobertura vegetal.

Assim como na classificação, na Análise Canônica é essencial que sejam usadas áreas representativas, e grandes o suficiente, para permitir que parâmetros estatísticos confiáveis seja usados. Por outro lado, as áreas que forem desprezadas na aquisição da estatística devem ser consideradas com cautela (ou totalmente desconsideradas!) nas imagens resultantes da Análise Canônica, pois irão conter informação distorcida.

9.7.2 Realce por Decorrelação

Um dos maiores problemas associados com as imagens produzidas pela APC diz respeito à dificuldade em interpretar os resultados, principalmente quando são usadas composições coloridas de três PCs. O motivo dessa dificuldade é relacionado ao significado das cores presentes numa composição colorida desse tipo. Usando-se bandas originais TM, por exemplo, é fácil de se prever o significado das cores em uma composição colorida, através da associação dessas bandas e das cores usadas em cada uma delas (RGB) com a assinatura espectral do material em questão. Assim, por exemplo, usando-se

a banda TM4 em vermelho, sabe-se que toda a biomassa verde aparecerá em vermelho, a banda TM3 em verde mostrará solos vermelhos na cor verde e assim por diante. Porém, uma vez que a os novos eixos das PCs são rotacionados, essa associação entre bandas originais e assinaturas espectrais se torna extremamente complexa (lembre-se que cada PC é uma adição ponderada de todas as bandas que compõem a imagem, algumas multiplicadas por -1!). Dessa forma, seria útil se, após a aplicação da APC e do aumento de contraste das novas imagens PCs, fosse possível restaurar as relações de cores originais, para efeito de interpretação. É precisamente essa a base teórica da técnica de **Realce por Decorrelação** (em inglês, "**Decorrelation Stretch**").

No Realce por Decorrelação, segue-se todos os passos usados em APC, até a fase final em que as imagens PCs tem o seu contraste aumentado, para que tenham as mesmas variâncias. Neste ponto a distribuição no espaço de atributos assemelha-se a um círculo, conforme visto na Figura 9.5 para o caso bidimensional. Uma vez que um círculo possui a mesma forma quando visto de qualquer angulo, **ele pode ser rotacionado ao redor do seu centro sem que o formato da distribuição seja alterado** e, portanto, sem que se mude a correlação entre as PCs representadas nesse espaço de atributos. **Se esse círculo fosse submetido à rotação inversa à da APC original**, as imagens resultantes teriam cores similares às das bandas originais. A única diferença é que o espaço de atributos será agora esférico e as imagens terão um intervalo de cores muito maior. A Figura 9.6 mostra, em um espaço bidimensional, como funciona essa técnica.

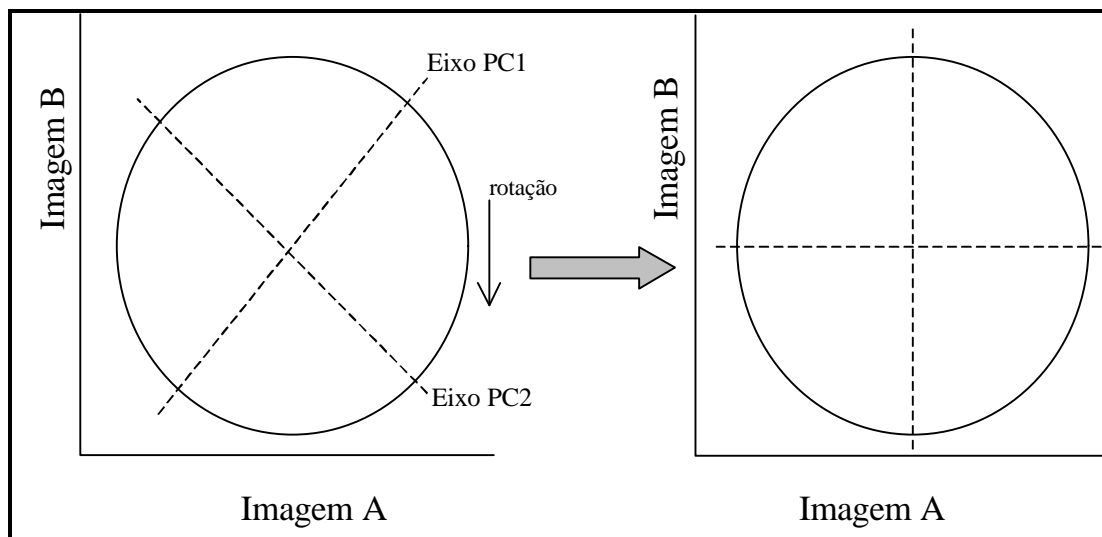


Figura 9.6 - No Realce por Decorrelação, após o aumento de contraste, a distribuição circular sofre a rotação inversa à sofrida pelos eixos das PCs, para que volte a ficar orientada segundo o sistema de coordenadas das imagens originais.

O Realce por Decorrelação, visa principalmente o realce da distribuição de cores, através da eliminação da correlação entre bandas. Por se tratar de uma técnica para realce de cores, **ela só faz sentido em ser aplicada a triplete de bandas**, ao contrário da Análise por Principais Componentes, que normalmente é aplicada a todas as bandas de uma imagem. Portanto, o primeiro passo no Realce por Decorrelação é selecionar o triplete desejado, para depois seguir os passos indicados.

Um último comentário sobre o Realce por Decorrelação se refere à sua implementação matemática. Uma vez obtida a matriz de auto-vetores do triplete de bandas selecionado para se aplicar essa técnica, **a única diferença em relação à APC é que o Realce por Decorrelação se utiliza da matriz transposta de auto-vetores**, isto é, as linhas passam a ser colunas e as colunas passam a ser linhas. Uma vez transposta a matriz, calcula-se cada banda decorrelacionada da mesma forma que as Pcs, ou seja, multiplicando-se cada PC pelos respectivos auto-vetores. Essa transposição da matriz de auto-vetores representa a rotação inversa dos eixos da PCs.