

White Paper

## Tecnologia Perfect Page a digitalização da página perfeita

Aproveite a ciência de imagem avançada usando o Perfect Page para otimizar as digitalizações



A digitalização de documentos é o pilar da transformação digital, e escolher a tecnologia certa é essencial para alcançar os resultados esperados. A capacidade de melhoramento de imagem de um scanner pode reduzir drasticamente o tempo gasto manuseando documentos e pode oferecer informações mais precisas para fluxos de trabalho automatizados. Muitas empresas ignoram os impasses envolvendo qualidade de imagem. No entanto, antecipar-se em investir na tecnologia certa pode gerar bons frutos em dividendos em pouco tempo.

#### Remova os impasses durante a captura

Muitas organizações ficam surpresas ao saber que preparar documentos para digitalização custa mais tempo e dinheiro do que o processo real de digitalização em si. De acordo com o chefe analista da AIIM, Bob Larrivee, “a preparação de documentos pode ser uma tarefa gigantesca quando se leva em conta o volume de papel a ser digitalizado”. Ele prossegue, “perto de um quarto das pessoas com que converso diz que a preparação de documentos é o maior problema na gestão de suas operações de digitalização, e aproximadamente o mesmo número fala que tem grandes dificuldades com documentos com problemas e exceções”.

As empresas enfrentam cada vez mais pressão para cortar os gastos, aprimorar a eficiência e otimizar os processos empresariais. O custo total de digitalização pode ser reduzido juntamente com o tempo de processamento ao se dedicar atenção aos elementos que parecem triviais por si sós, mas que totalizam em grandes economias quando tomados em conjunto.

#### A qualidade é essencial ao digitalizar documentos

Um desafio bastante comum é atender aos requisitos de qualidade de imagem das partes interessadas durante projetos de digitalização de documentos. Os requisitos podem variar de acordo com o tipo de documento a ser digitalizado. É esperado um tipo de saídas para formulários impressos, cartas ou pedidos por escrito com fotos embutidas; e outro tipo para relatórios gráficos, em que as cores e as ilustrações são importantes.

Pergunte ao destinatário final como as informações serão utilizadas nos fluxos de trabalho empresariais:

- Os documentos serão digitalizados para fins de arquivamento?
- Quais requisitos regulatórios e de conformidade devem ser atendidos?
- O fluxo de trabalho envolve extração automática de informações chave para iniciar um fluxo de trabalho empresarial? (Por exemplo, processamento de faturas)

As repostas para perguntas desse tipo determinarão as necessidades de melhoramento de imagem da saída final.

#### A preparação de documentos é um grande impasse

A digitalização em si é uma pequena parte do processo de digitalização. Em ambientes de digitalização, não é raro encontrar metade (ou até mais) do pessoal concentrado em pré-classificar e preparar documentos. Até mesmo em uma operação de digitalização configurada para processar o mesmo tipo de documento, como faturas, há vários desafios: os operadores tem que enfrentar uma grande variedade de tipos de papel e tamanhos de documentos, bem como fundos coloridos diferentes e um misto de documentos em retrato e paisagem. Frequentemente as operações têm funcionários dedicando tempo para classificar os documentos manualmente em vez de deixar a tecnologia fazer o trabalho por eles. Os scanners certos podem fazer esse trabalho automaticamente, economizando custos com trabalho extra.

Além disso, há ainda mais trabalho manual envolvido ao separar documentos que são considerados desafiantes. Quando no meio de um trabalho é encontrado um documento de difícil leitura, este é digitalizado em um lote separado com configurações especiais, geralmente em alta resolução, o que resulta em arquivos desnecessariamente grandes ou altos níveis de ruído no fundo. Caso esse documento seja deixado no fluxo normal de documentos sem melhoramento de imagem, a verificação pelo controle de qualidade tende a rejeitá-lo. Isso resulta na necessidade de redigitalizar ou realizar a indexação ou extração de dados manualmente. Qual a típica e surpreendente maneira de contornar esse problema? Os operadores de digitalização usam uma copiadora com uma alta configuração de contraste para reimprimir o documento. Isso é um esforço manual significativo que atrasa as operações e aumenta custos que poderiam ser facilmente evitados.

## Trabalho desnecessário pós-digitalização impactam a rentabilidade e a satisfação dos clientes

Ao usar o reconhecimento óptico de caracteres (OCR), a qualidade de imagem é peça fundamental. Documentos mal digitalizados, borrados com falhas, imagens mal processadas ou texto muito claro resultarão em erros no reconhecimento e os resultados do OCR serão afetados. Até mesmo um ponto percentual a menos na taxa de leitura OCR pode significar um grande custo desnecessário associado ao processamento de exceções.

Documentos reprovados na verificação de controle qualidade e aqueles não reconhecidos durante o processo de OCR, precisam ser redigitalizados e ter os campos de índices inseridos manualmente, o que gasta tanto tempo como dinheiro. O nível de processamento de exceções também tem um impacto negativo adicional na eficiência empresarial. Isso afeta as métricas de sucesso empresarial, tais como tempo de processamento de pedidos, o que tem um efeito dominó decisivo nos níveis de satisfação do cliente.

## Reduza o tempo gasto na preparação de documentos

A possibilidade de alimentar lotes mistos, da exata maneira como chegam, sem sacrificar rapidez e garantindo uma qualidade de saída alta e consistente representa uma imensa oportunidade para as organizações melhorarem seus resultados. Isso também reduz drasticamente o tempo desperdiçado em pré-classificar documentos manualmente.



O objetivo da equipe de processamento avançado de imagem Kodak Alaris é, com o tempo, eliminar completamente a preparação de documentos. A Kodak Alaris traz consigo uma rica bagagem em ciência de imagem focada na qualidade incomparável de imagem para todos os tipos de documentos. A tecnologia Perfect Page oferece possibilidades que refletem o estado da arte em melhoramento de imagem até mesmo para os documentos mais desafiadores e lotes de documentos mistos. Todos os recursos abordados no decorrer deste manual são cruciais para a tecnologia Perfect Page. Quando estiver pensando em investir em um novo escâner, esteja certo de levar em conta todas as maneiras que a tecnologia certa pode ajudar seu negócio a economizar tempo, reduzir custos e melhorar a precisão em suas operações de digitalização. Continue lendo para saber detalhes sobre como o Perfect Page aprimora imagens para a maioria dos tipos de aplicações de digitalização mais comuns e mais desafiadores.

**Enquadramento e autocorte:**

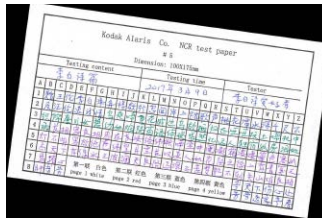
Correspondências raramente chegam em uma pilha organizada, principalmente quando os documentos têm tamanhos diferentes. As imagens precisam ser alinhadas (ou enquadradas) para que possam ser usadas futuramente em processos como reconhecimento automático de texto, escrita ou marcas de verificação. É preciso alinhar a imagem de maneira que os lados do documento estejam paralelos às bordas da imagem de saída. Dependendo da aplicação da digitalização, uma borda preta pode ser necessária ou não. Documentos com extenso fundo preto, como demonstrado na primeira imagem a seguir, comprometem qualquer uso que se possa fazer da imagem. Portanto, é importante cortar a imagem para capturar somente a área que representa o documento.

Um corte mal feito pode gerar problemas na localização de dados em formulários. A maioria dos softwares ainda encontra os dados usando os cantos do documento como pontos localizadores em vez de “fiduciais” (pontos de referência por comparação), que são uma opção melhor. Para documentos longos, é preciso basear o corte em uma análise de quatro cantos e não em um método baseado no comprimento e em dois cantos. Isso dará resultados superiores para páginas que não são perfeitamente retangulares, como demonstrado na imagem a seguir.

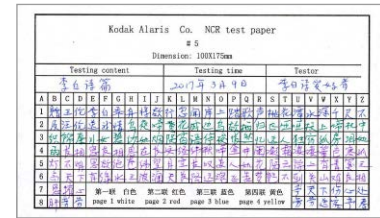
Output without deskew and cropping



Data loss due to faulty deskew



Autocrop and deskew



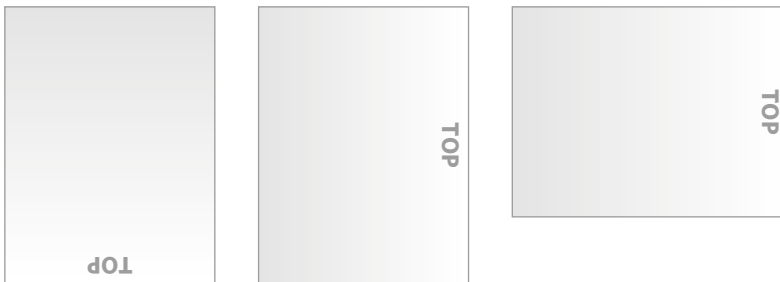
**Orientação automática:**

A possibilidade de alimentar o escâner com volumes mistos de documentos em retrato e paisagem e obter uma série de imagens perfeitamente orientadas é um ganho enorme em eficiência. Às vezes, algumas empresas precisam digitalizar em paisagem para obter maiores velocidades de transferência ou para evitar problemas com os documentos que afetariam a alimentação, como bordas aderentes.

Na maioria das vezes, uma rotação fixa de 90 graus para documentos em paisagem é um bom começo. Nas vezes em que documentos deslizam durante a preparação, o scanner precisa ser inteligente o bastante para ajustar a orientação imediatamente.

A orientação automática deve ser baseada no conteúdo, a fim de detectar e corrigir a orientação dos documentos independentemente de como eles foram inseridos. Para digitalizações duplas em paisagem, é recomendável combinar a rotação padrão e a orientação automática. E para organizações internacionais, isso precisa funcionar independentemente do idioma usado.

Input



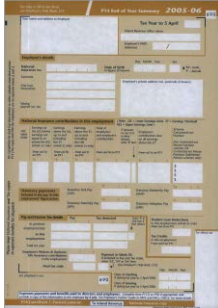
Output



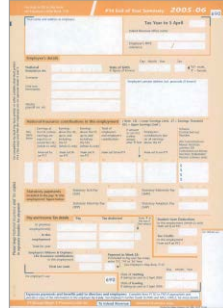
### Brilho automático:

Não há necessidade de usar uma copiadora para alterar o brilho de um documento com pouco contraste. A tecnologia de digitalização ajusta automaticamente a otimização de brilho de imagem, sem perda na velocidade ou na taxa transferência em documentos coloridos ou em escala de cinza. Clareando ao máximo possível as cores mais claras de cada imagem e escurecendo ao máximo as cores escuras, tanto a percepção de qualidade de imagem quanto a legibilidade dos documentos são melhoradas significativamente. Esse recurso traz muitas vantagens para o arquivamento de documentos.

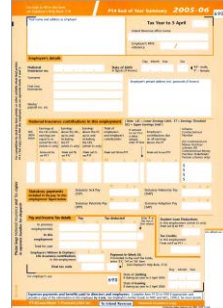
Input



Default



Autobright



### Nitidez:

Aumentando o contraste dos contornos em uma imagem, a nitidez automática deixa os objetos na imagem mais vivos. Isso aprimora a aparência do documento melhorando as taxas de leitura OCR.

Sharpening sample

None



Normal



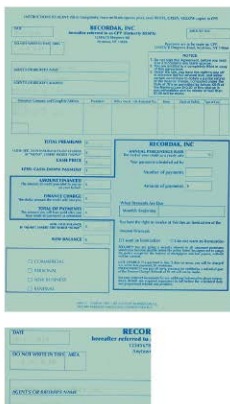
High



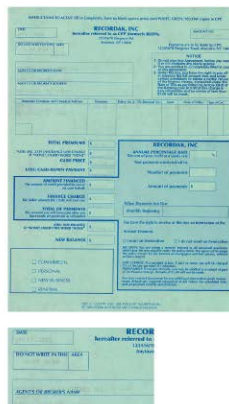
### Suavização inteligente de cor de fundo:

As cores de fundo em uma imagem colorida ou em escala de cinza podem não ser uniformes. A suavização de imagem minimiza a variação de cores, oferecendo uma imagem mais uniforme que se parece mais com um documento criado digitalmente. Isso geralmente reduz o tamanho de imagens comprimidas. Documentos ou formulários cujo primeiro plano (texto, linhas, etc.) precisa receber destaque especial podem ser aprimorados por meio do realce inteligente de primeiro plano.

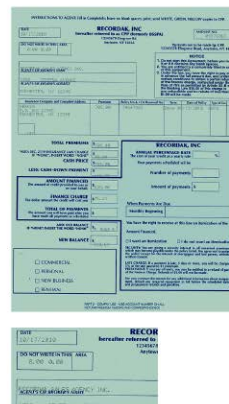
No color smoothing



Color smoothing but foreground hard to see



Color smoothing and foreground boldness added



## Remoção de riscos:

A tecnologia de filtragem de riscos soluciona um dos problemas de qualidade de imagem mais comuns: linhas pretas verticais em uma imagem original não enquadrada. Isso é frequentemente causado pelo acúmulo de poeira com o passar do tempo sob o local onde o scanner é alojado. É necessário limpar regularmente para evitar que isso aconteça, mas caso esse problema ocorra, a tecnologia de melhoramento de imagem é capaz de removê-lo ou reduzi-lo.

Before (streak filtering off)

Color

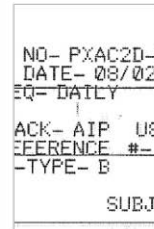
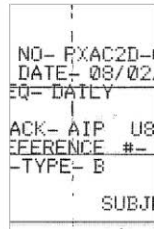


Black and white



After (streak filtering on)

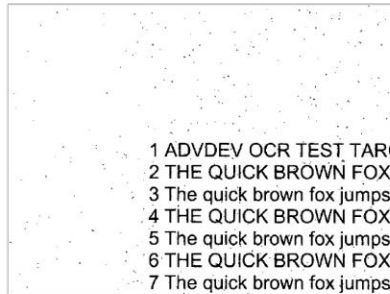
Black and white



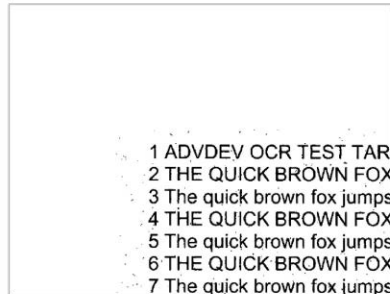
## Redução de ruído:

Outro desafio bastante comum ao converter documentos coloridos para imagens em preto e branco é a presença de pontos pequenos (ou maiores), chamados geralmente de "ruído", causados por poeira ou baixa qualidade do papel, como papel reciclado. Algoritmos de remoção de ruído removem pontos únicos (pixel isolados), grupos de pixels (a maioria dos casos, via de regra) ou até mesmo grupos maiores de pixels (remoção de ruído no fundo) para melhorar a aparência de um documento.

Noise reduction



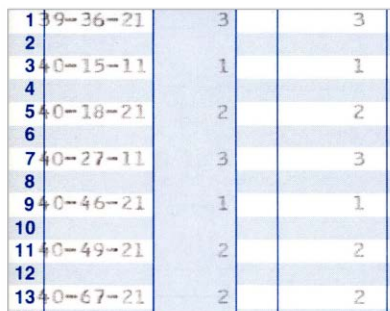
Background noise removal



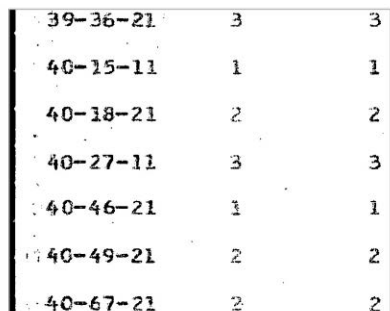
## Remoção de todas as cores:

O uso de remoção de cores é uma prática padrão durante o processamento automatizado de formulários, em que uma imagem bitonal otimizada é necessária para melhorar as taxas de leitura OCR. As empresas usam em um mesmo formulário diferentes cores para eliminação juntas. Uma tecnologia inteligente pode identificar e remover eletronicamente múltiplas cores dominantes ou cores pré-definidas para reduzir a intervenção manual e garantir as melhores taxas de leitura.

Color image



All-color drop-out



**Fluxo duplo:**

É importante gerar ao mesmo tempo uma digitalização colorida e uma em preto e branco. Uma imagem colorida é geralmente a escolha preferida para fins de arquivamento e para qualquer interação humana durante o processamento. Para processamento automático por indexação ou extração de dados, uma imagem bitonal otimizada (em preto e branco) é necessária. O fluxo duplo cria os dois tipos de arquivo durante o processo de digitalização.

**Segmentação de fotos:**

Documentos que contêm texto e imagens, como pedidos de passaporte ou seguros, precisam receber tratamento especial. Ao diferenciar áreas com fotos de áreas de texto, recursos tecnológicos podem aplicar diferentes técnicas de processamento a cada área e aplicar o recorte de foto às imagens coloridas. Caso uma imagem bitonal seja necessária, é possível usar a tecnologia de matização automática para fazer a foto parecer quase uma imagem em escala de cinza, enquanto outros elementos, como códigos de barras, não são matizados, a fim de garantir o reconhecimento destes.

Input

Binarized

Segmentation Applied



**Processamento inteligente de limiar (iThresholding) com taxas de leitura OCR e binarização aprimoradas:**

A binarização, por exemplo, a conversão de imagens coloridas para preto e branco, é vital para todos os recursos de extração de dados e é, portanto, a base para qualquer processamento de imagens. Em um lote com documentos de qualidades diferentes, os requisitos para a binarização e o melhoramento variam de um documento para outro. Recursos de tecnologia inteligente analisam o primeiro plano e o fundo das imagens de documentos, avaliando o brilho e o contraste, e então configuram os melhores valores de limiar. Isso otimiza a qualidade geral das imagens e o tamanho dos arquivos, principalmente em documentos com sombreamentos e semitons que causam problemas de reconhecimento de caracteres.

Original form

Output binarized

Output optimized

# Conclusão

O processamento avançado de imagens pode agilizar a digitalização com precisão de documentos e como estes são repassados aos sistemas empresariais. A tecnologia Perfect Page oferece várias técnicas para melhorar a qualidade de imagem, dispensando os numerosos e incômodos esforços manuais na preparação de documentos ou no processo de digitalização. As possibilidades ilustradas neste artigo são apenas uma pequena fração dos recursos do Perfect Page disponíveis na nossa linha premiada de scanners de documentos e softwares de digitalização. Clientes e parceiros reconhecem o Perfect Page como uma das principais vantagens dos scanners e softwares que compõem o Alaris IN2 Ecosystem. Ao avaliar uma nova solução de captura de documentos, certifique-se de dar a importância devida a esses aspectos em relação ao impacto nos custos totais de propriedade. Solicite aos fornecedores que você está avaliando para demonstrar a qualidade de imagem deles e avalie os impactos na preparação e pós-digitalização de seus documentos e os custos associados.

## A tecnologia Perfect Page permite:

- Reduzir a preparação de documentos a simples etapas, como remover grampos, sem necessidade alguma de organizar os documentos previamente.
- Processar lotes de documentos mistos sem ter que organizar por orientação, tamanho, tipo ou forma.
- Aumentar significativamente os resultados de OCR/ICR, bem como reduzir o processamento de exceções a apenas uma fração.

A tecnologia Perfect Page é uma parte integrante do [Alaris IN2 Ecosystem](#) e garante a qualidade de imagem excepcional associada à Kodak Alaris. O Alaris IN2 Ecosystem é uma combinação poderosa dos melhores scanners, softwares, serviços e parcerias da categoria, desenvolvida para simplificar a captura de informações e possibilitar que os clientes transformem o [caos de dados](#) em informações primordiais para impulsionar a eficiência.

### Want to learn more?

[AlarisWorld.com](http://AlarisWorld.com)

### Contact us:

[AlarisWorld.com/go/contactus](http://AlarisWorld.com/go/contactus)

# Alaris

a Kodak Alaris business

All trademarks and trade names used are the property of their respective holders.

The Kodak trademark and trade dress are used under license from Eastman Kodak Company.

© 2018 Kodak Alaris Inc.  
TM/MC/MR: Alaris, Perfect Page  
03/18

