

Método Genérico para Estimação e Modelagem do Erro RMS em Dados de Profundidade de Sensores para Visão 3D

Luis E. Ortiz, Elizabeth V. Cabrera, Luiz M. G. Gonçalves
Programa de Pós-graduação em Engenharia Elétrica e de Computação
Universidade Federal do Rio Grande do Norte
Avenida Senador Salgado Filho, 3000, CEP 59.078-970
Natal, RN, Brasil

Resumo—No mercado existem vários dispositivos que são usados na visão computacional ou robótica para estimar informação de profundidade de uma cena. Como cada sensor usa uma diferente tecnologia (estéreo, luz estruturada e tempo de voo) para obter os dados de profundidade, estes podem conter erros. Neste artigo, introduzimos um método versátil para a estimação do erro RMS em dados de profundidade fornecidos por qualquer dispositivo que seja capaz de gerar dados do tipo RGB-D, isto é, uma imagem (RGB) e um mapa de profundidade (D) ao mesmo tempo. Para verificação do método foi construído um sistema embarcado baseado na placa NVIDIA Jetson TK1 e usados três sensores, quais sejam, as duas versões do MS Kinect e a câmera estéreo ZED. No momento da coleta de dados, foram estabelecidos os modelos matemáticos do erro RMS para cada dispositivo e, ao final, foi feita uma análise da exatidão de cada um usando a metodologia proposta. Os erros de profundidade estimados pelo nosso método são comparados com aqueles estimados por outras abordagens, validando, assim, sua exatidão e precisão. Devido à sua versatilidade, o método proposto pode ser usado para estimar rapidamente a qualidade de vários sensores RGB-D, beneficiando aplicações como SLAM e reconhecimento de objetos.

I. INTRODUÇÃO

Sistemas de visão artificial procuram simular a visão humana mediante o uso de dispositivos que capturam dados (2D ou 3D) do ambiente baseados em diversas tecnologias, tais como as câmeras estéreo, os sensores de luz estruturada e os sensores baseados em tempo de voo (time-of-flight). Estes dados são geralmente processados através de técnicas de visão computacional, fazendo com que sistemas artificiais embarcados em robôs, por exemplo, possam perceber o ambiente e responder autonomamente a determinadas situações. Uma etapa normalmente incluída no processamento dos dados refere-se ao uso de métodos para geração de modelos 3D a partir de imagens 2D, determinando assim, a profundidade dos objetos em uma cena.

Existem vários dispositivos, com seus softwares associados (SDK), que podem ser usados para determinar a profundidade de objetos, tais como o MS Kinect v1/v2, Asus Xtion Pro, PrimeSense Carmine, e as câmeras estéreo PG Bumblebee XB3, Stereolabs ZED, e Minoru 3D, entre outras. Como todos

são dispositivos que estimam medidas de profundidade, é normal que tais estimações da profundidade contenham um certo grau de erro, associado ou não ao tipo de dispositivo que esteja sendo usado.

Dos trabalhos existentes [1]–[4] e a partir de testes que realizamos em laboratório, é possível noticiar que, no caso das câmeras estéreo, os erros são causados pelo uso de lentes com alta distorção, disposição inadequada das câmeras, captura de imagens com baixa resolução, problemas de iluminação e calibração com alto erro de reprojeção. Conforme estudos encontrados na literatura a respeito [5]–[7], no MS Kinect e outros dispositivos semelhantes, os erros aparecem principalmente devido à baixa taxa de captura e de resolução. Isso deve-se principalmente ao fato de os mesmos terem sido projetados para entretenimento, e não para desenvolvimento de aplicações científicas e/ou tarefas de alta precisão. Mas, nota-se que, apesar das limitações, os dispositivos acima citados têm servido de base para muitos projetos de pesquisa em visão computacional e robótica [8]–[10], principalmente por causa do baixo custo e precisão relativamente aceitáveis.

Um dos problemas que foi identificado em relação aos dispositivos acima é a carência de dados confiáveis de seu erro. Isto é, justamente por serem dispositivos ainda recentes, há uma carência de métricas para mensurar a raiz do valor quadrático médio do erro, também denominado de erro RMS, inerente à determinação da profundidade que os sensores conseguem retornar.

Assim, por não haver encontrado na literatura trabalhos a respeito, propomos neste artigo um método genérico e versátil para a estimação do erro RMS (seu modelo matemático) nos dados de profundidade que sejam fornecidos por qualquer dispositivo que seja capaz de gerar uma imagem (RGB) e um mapa de profundidade (D) ao mesmo tempo (RGBD), visando determinar melhor sua precisão.

II. TRABALHOS RELACIONADOS

O MS Kinect v1 é frequentemente utilizado para o desenvolvimento de aplicações em visão computacional, principalmente, por causa do seu custo relativamente baixo e do desempenho aceitável na captura de dados de profundidade.

¹Este trabalho relaciona-se com uma dissertação de Mestrado

No entanto, ele tem limitações definidas em vários trabalhos como, por exemplo, o de Khoshelham [11], onde é determinado matematicamente e, de forma empírica, que o erro nas medições de profundidade aumenta de forma quadrática com o aumento da distância entre o sensor e o alvo. Para aplicações de mapeamento, recomenda-se que os dados sejam obtidos a uma distância de 1 até 3 m do sensor. Também é proposto um modelo matemático para calcular o erro a uma determinada distância.

Ao fazer uma análise empírica da precisão do MS Kinect e o Asus Xtion Pro, no trabalho de Haggag [12], foi estabelecido que o MS Kinect é mais sensível aos objetos que refletem iluminação. O trabalho conclui, também, que a luz infravermelha emitida pelo sensor afeta a precisão da profundidade de um outro sensor quando estes se juntam para capturar a mesma cena sem a separação suficiente. Devido aos erros dos dispositivos RGB-D em longas distâncias, alguns trabalhos como o de Yang [13] propõem juntar dois ou mais sensores para melhorar a precisão dos dados de profundidade na captura de objetos a uma distância maior que 4 m. Concluiu-se que, um objeto coberto de materiais que absorvem a luz pode gerar menos luz infravermelha refletida de volta para o MS Kinect e, portanto, dados de profundidade errados. Outros fatores, tais como, o consumo de energia, fiação complexa e alto uso dos recursos do computador portátil também limitam o uso do sensor.

O artigo de Aguilar [14] avalia as câmeras estéreo CCD (*Charge Coupled Device*). Para verificar sua precisão, são usados alguns objetos de calibração como grades, placas, esferas e etc. Assim, são determinados os limites do erro em certas condições de operações específicas. No trabalho de Luo [15] é calculado o erro médio e o desvio padrão na reconstrução da superfície de um corpo, usando uma câmera estéreo. Neste trabalho, é apresentada uma abordagem matemática usando a equação da superfície cilíndrica para obter o raio do cilindro e compará-lo com o valor medido pelo dispositivo. Os resultados indicam que a visão estéreo pode medir com precisão a superfície em várias curvaturas.

Na literatura existem outras abordagens para determinação do erro nos dados de profundidade de sensores 3D, como os métodos estatísticos [16], geométricos e probabilísticos [2], empíricos [17]–[19], entre outros. Depois do análise de vários trabalhos relacionados pode-se dizer que a principal contribuição de nosso método é o fornecimento, aos desenvolvedores de aplicações de visão computacional e visão robótica, de um ferramental capaz de expor o erro associado ao uso do sensor para visão 3D e que, permita também, dimensionar o erro de sua aplicação. Por ser versátil, o método proposto pode ser usado para determinar o erro RMS nos dados de profundidade de qualquer dispositivo capaz de fornecer dados RGB-D.

Como contribuição adicional, apresentamos uma análise completa da câmera ZED. Ao contrário dos trabalhos relacionados a este, é demonstrado de forma matemática e experimental o erro da câmera em todos os eixos coordenados, assim como, são fornecidas as curvas que modelam o erro RMS para cada resolução da câmera.

Aplicações com sensores de profundidade, montados sobre robôs aéreos, terrestres e aquáticos, que necessitam de uma estimação do erro podem usar do método. Um exemplo imediato de aplicação é o uso de robótica probabilística para localização e mapeamento. Outras tarefas como monitoramento ambiental, manutenção de obras de engenharia e segurança pública podem contar com esta abordagem para obter informações do erro RMS.

III. APRESENTAÇÃO DO MÉTODO

O problema de determinação do erro a ser modelado matematicamente neste trabalho pode ser formalizado simplesmente como:

Dado dois conjuntos de estruturas (features) do tipo canto, um ponto ideal p_i (obtido de um padrão manualmente) com coordenadas referenciadas ao sistema de câmera $(Xc_{p_i}, Yc_{p_i}, Zc_{p_i})$, e outro ponto medido (estimado) pelo SDK do sensor 3D p'_i com coordenadas no sistema de pixels dadas por $(u_{p'_i}, v_{p'_i})$ e com coordenadas no sistema de câmera dadas por $(Xc_{p'_i}, Yc_{p'_i}, Zc_{p'_i})$, qual é o erro RMS entre p_i e p'_i ?

Nos tratamos o problema da estimação do erro RMS em sensores 3D seguindo alguns passos a serem realizados: a captura de imagens RGB, que devem estar alinhadas com o mapa de profundidade; a detecção de estruturas do tipo cantos varrendo todo o alcance do sensor analisado; o registro e alinhamento das nuvens de pontos usando as coordenadas 3D (ideais e estimadas); e por fim a modelagem matemática do erro RMS. Esses passos serão tratados a seguir, separadamente.

A. Captura de dados

O propósito deste trabalho é avaliar de forma global os dispositivos, considerando o *hardware* e algoritmos usados para a geração de dados de profundidade. Para que o problema seja solucionado é fundamental que alguns aspectos (ou restrições) sejam pré-estabelecidos. Como restrições fundamentais podemos citar:

- Os parâmetros intrínsecos e extrínsecos do dispositivo analisado devem ser conhecidos;
- O dispositivo de captura deve ser previamente calibrado para que forneça imagens e mapas de profundidade retificados, ou seja, presumivelmente sem distorções radial e tangencial.
- Deve-se alinhar a imagem RGB com o mapa de profundidade para que possa existir uma relação entre cada pixel $(u_{p'_i}, v_{p'_i})$ e sua respectiva profundidade $Zc_{p'_i}$, como mostrado na Figura 1.

B. Padrão usado e detecção de cantos

Para estimar o erro RMS, são necessários os dados ideais, que irão servir como base para a comparação com os dados estimados (calculados ou medidos). Esses dados podem ser obtidos manualmente a partir de um padrão de fácil construção e com baixo custo, como os tabuleiros de xadrez usados comumente neste tipo de abordagem. Em visão computacional,

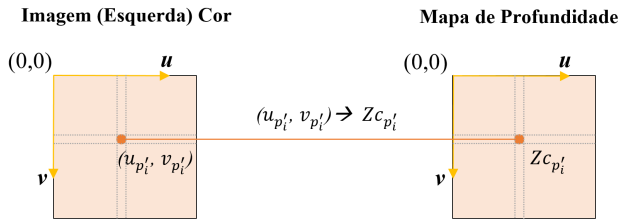


Figura 1. Relação entre uma imagem RGB e seu mapa de profundidade

estes tabuleiros são amplamente usados, pois seus cantos são de fácil detecção quando comparados com outros padrões. O principal problema que encontramos ao usá-los inicialmente foi na detecção de seus cantos a longas distâncias. Ou seja, no problema aqui tratado, os algoritmos de detecção e ordenação devem ser eficazes para atingir o alcance máximo do sensor analisado (até 20 metros ou mais). Os algoritmos tradicionais foram desenvolvidos na literatura, em sua maioria, para dispositivos RGB-D que tem curto alcance, mais especificamente para distâncias que vão desde 1 até 10 m. Nas câmeras estereó atuais este alcance pode ser duplicado, para 20 m ou mais, e para que o método funcione sem ambiguidades, a detecção deve ser de múltiplos padrões (com pontos distribuídos em uma configuração 3D sem serem coplanares). Ainda, pela questão da distância, é necessário que se tenha padrões de vários tamanhos. O algoritmo de detecção de cantos que se adapta a nossos requisitos é a proposta de Geiger [20].

C. Registro e alinhamento de nuvens de pontos

Nossa estratégia para estimar o erro nos dados de profundidade é projetar os cantos p_i de um padrão, que estão em um sistema de coordenadas de pixel $(u_{p'_i}, v_{p'_i})$ para um sistema de coordenadas câmera $(Xc_{p'_i}, Yc_{p'_i}, Zc_{p'_i})$ usando a Equação 1, onde (C_x, C_y) é o ponto onde o raio principal atravessa o plano da imagem e (f_x, f_y) é a distância focal. Dessa maneira, obtém-se as coordenadas 3D ideais (do padrão) e as estimadas (medida pelo dispositivo) dos cantos e pode-se representar esses cantos ideais e estimados como duas nuvens de pontos. As coordenadas ideais servem como base para o cálculo do erro RMS das outras coordenadas, estimadas pelos dispositivos testados.

$$Xc_{p'_i} = \frac{u_{p'_i} - C_x}{f_x} Zc_{p'_i} \quad Yc_{p'_i} = \frac{v_{p'_i} - C_y}{f_y} Zc_{p'_i} \quad (1)$$

Antes de calcular o erro RMS, as nuvens de pontos ideal e estimada devem ser registradas e alinhadas. Inicialmente, a primeira está representada em um sistema de coordenadas mundo e a segunda no sistema de câmera. O objetivo é encontrar as posições e orientações das nuvens de pontos em uma estrutura de coordenadas globais, de modo que ambas nuvens tenham o mesmo *frame* de referencia. Em nosso caso, as duas nuvens têm índices ou correspondências conhecidas, têm o mesmo número de pontos, e cada ponto em uma nuvem tem seu correspondente na outra. Então o objetivo final é obter

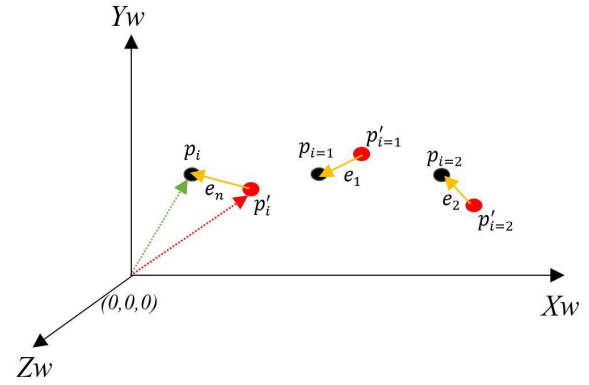


Figura 2. Processo do cálculo do erro RMS em nuvens de pontos.

uma matriz de transformação rígida (4×4) que representa a rotação e translação $[R|t]$ de uma nuvem com respeito a outra. Esta transformação é obtida usando o algoritmo descrito por Arun [21].

D. Estimativa do erro RMS e modelagem matemática

Após o registro, surge o problema de juntar todos os erros individuais das coordenadas de cada canto em apenas uma métrica. Uma das melhores formas para fazer isso é usar o valor RMS, já que é uma boa medida de exatidão, pois quantifica a diferença entre dois conjuntos de dados e amplia os grandes erros, considerando valores negativos e positivos.

O erro quadrático $\sum^2$ é calculado usando a Equação 2, onde simplesmente realiza-se a soma das distâncias euclidianas $|e_1|, |e_2|, \dots, |e_n|$ medidas entre cada canto ideal $(p_1, p_2, \dots, p_i)$ até cada canto estimado $(p'_1, p'_2, \dots, p'_i)$, tal como ilustrado na Figura 2. O erro RMS é obtido a partir da Equação 3.

$$\sum^2 = \sum_{i=1}^n \| p_i - (Rp'_i + T) \|^2 \quad (2)$$

$$e_{RMS} = \sqrt{\frac{\sum^2}{n}} \quad (3)$$

Após o cálculo dos valores de erro RMS obtém-se uma representação matemática que possa ser usada em outras aplicações de visão robótica. Para isso, é usada a regressão por mínimos quadrados com a função exponencial da Equação 4, onde a e b são os coeficientes de ajuste e Z_j é a distância entre o padrão e o sensor 3D testado. O ajuste de curva é avaliado usando as métricas Soma de Quadrados devido ao Erro (SSE), $R_{quadrado}$ e Raiz do Erro Quadrático Médio (S) [22].

$$e_{RMS} = f(Z_j) = ae^{bZ_j} \quad (4)$$

IV. EXPERIMENTOS E RESULTADOS

Para verificar e testar o método proposto, é usado o sistema embarcado cuja montagem é descrita em nosso trabalho anterior [23]. Ele é composto por uma placa NVIDIA Jetson TK1, os MS Kinect's v1, v2 e a câmera Stereolabs ZED.

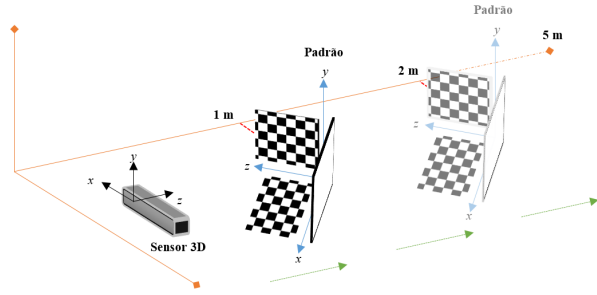


Figura 3. Processo dos experimentos para determinação do erro RMS, usando um padrão 3D.

Também são usados padrões (tabuleiros de xadrez) com dois e três planos. São feitos dois tipos de testes, um para obter uma comparativa do erro RMS de todos os dispositivos mencionados acima, e outro para determinar apenas o erro RMS da câmera ZED para cada uma de suas resoluções de captura (*frames* individuais de 2208×1242 , 1920×1080 , 1280×720 e 672×376 px). Após a coleta de dados, foi realizado um ajuste de curva para obter a equação que relaciona o erro RMS e a distância entre o padrão e o sensor (profundidade).

A. Determinação do erro RMS: Câmera ZED

Com o propósito de determinar a exatidão dos dados de profundidade da câmera ZED, em todas as suas resoluções de captura, foi realizado o processo ilustrado na Figura 3. Um padrão com dois tabuleiros de xadrez ortogonais um ao outro foi usado e cada um deles tem (3×4) cantos internos. O método foi testado para verificar se era compatível com este tipo de padrão.

Para conseguir uma detecção correta dos cantos a longas distâncias, foi determinado empiricamente o tamanho dos quadrados pretos e brancos com 20 cm de lado. Usando esse padrão em cada experimento, obtém-se um conjunto de amostras (nuvens de pontos) com vinte e quatro valores de profundidade e seus respectivos erros. O processo é repetido para várias distâncias entre o padrão e a câmera Z_j , variando-se essas distâncias na faixa de 1 até 20 m, a cada 1 m. Assim, foram realizados 20 repetições (testes) do mesmo experimento, uma para cada resolução da câmera, totalizando em 80 repetições.

Depois de avaliar os ajustes gráficos e numéricos usando as métricas SSE , $R_{quadrado}$ e S foi determinado empiricamente, que o melhor ajuste para os dados é o exponencial (Equação 4) com os coeficientes mostrados na Tabela I. Em resumo, a Equação 4 com os respectivos coeficientes a e b da Tabela I pode ser usada para calcular o erro RMS para qualquer profundidade Z_j , em qualquer resolução da câmera ZED.

Ao comparar as quatro curvas apresentadas na Figura 4a, pode-se ver que o sistema não apresenta uma variação elevada na amplitude do erro RMS para distâncias de até 7 m. A partir dos 8 metros, pode-se ver um rápido aumento na amplitude das curvas “ 632×376 px” e “ 1280×720 px”. Isto deve-se ao

Tabela I
COEFICIENTES PARA O AJUSTE EXPONENCIAL DOS DADOS DE ERRO RMS DA CÂMERA ZED.

Resolução [px]	Coeficientes
2208×1242	$a=0,01805$, $b=0,1746$
1920×1080	$a=0,0106$, $b=0,2215$
1280×720	$a=0,0184$, $b=0,2106$
672×376	$a=0,0115$, $b=0,2986$

Tabela II
COEFICIENTES PARA O AJUSTE EXPONENCIAL DOS DADOS DE ERRO RMS DO MS KINECT V1, V2 E DA CÂMERA ZED.

Sensor	Coeficientes
MS Kinect v1	$a=0,002002$, $b=0,9603$
MS Kinect v2	$a=0,001224$, $b=1,074$
Stereolabs ZED	$a=0,003049$, $b=0,6242$

fato de a câmera deixar de fornecer dados de profundidade para estas resoluções a uma distância maior. Por outro lado, as curvas “ 1920×1080 px” e “ 2208×1242 px” permanecem com um mínimo de variação de amplitude (até 13 m) entre elas. Mais detalhes da análise dos erros na câmera ZED estão disponíveis em nosso trabalho publicado recentemente [24].

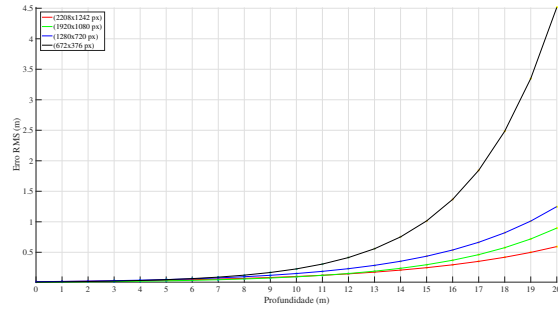
B. Determinação do erro RMS: Comparação entre sensores

Neste segundo grupo de experimentos, captura-se com cada sensor uma imagem RGB e o mapa de profundidade do padrão que contém três tabuleiros de xadrez (ortogonais). Este processo é repetido colocando o padrão a várias distâncias de cada sensor (até 3 m). As imagens são capturadas com as resoluções máximas de cada dispositivo utilizado, o MS Kinect v1 (640×480 px), o MS Kinect v2 (512×424 px) e a câmera ZED (2208×1242 px). Com o uso destas resoluções é possível avaliar cada dispositivo em suas capacidades máximas. Os três sensores são testados no mesmo local, fechado e com luz artificial, para obter bons dados de profundidade dos sensores RGB-D, já que eles são mais sensíveis à luz natural.

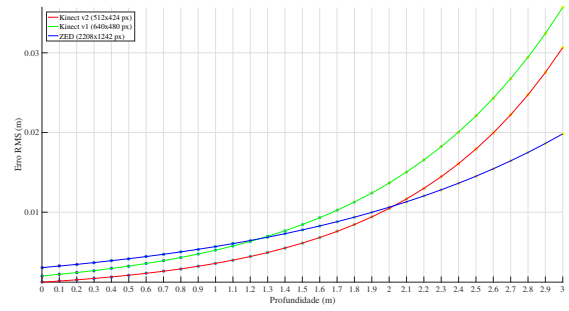
Com os valores do erro RMS e profundidade, os modelos matemáticos para o MS Kinect v1, v2 e a câmera Stereolabs ZED são obtidos. A Figura 4b representa o ajuste dos dados de erro a uma curva, usando o modelo de regressão por mínimos quadrados.

O uso do ajuste de tipo exponencial é justificado pela correlação e tendência que têm os dados. No caso do MS Kinect v1 os valores de erro RMS e profundidade têm uma correlação de 0,92; para o MS Kinect v2 é de 0,95 e na câmera Stereolabs ZED é de 0,98. A Tabela II lista os coeficientes a e b para a Equação 4.

Comparando as três curvas apresentadas na Figura 4b, pode-se ver que a amplitude do erro RMS do MS Kinect v2 é menor do que da ZED e do MS Kinect v1 apenas para as distâncias até 2 m. A partir dos 3 m, o erro RMS da ZED é menor do que os erros dos dois Kinects, devido ao fato de que seu mapa de profundidade possui melhor resolução em comparação com outros dispositivos. Pode ser notado também que a curva do



(a) Erro RMS para cada resolução da câmera ZED, de 1 até 20 m.



(b) Comparativa do Erro RMS no MS Kinect v1, v2 e ZED de 1 até 3 m.

Figura 4. Curvas do erro RMS.

erro do Kinect v1, a partir de 1,3 m, mantém um aumento de amplitude maior do que as outras curvas. No caso do MS Kinect v1 e v2, o aumento do erro é mais acentuado quando a distância for maior que 3 m, pois os dois sensores deixam de fornecer dados de profundidade quando estão a grandes distâncias.

V. VALIDAÇÃO

A análise do erro de dados de profundidade do Kinect v1 tem sido abordada em diversos trabalhos da literatura, no entanto cada autor propõe variantes na metodologia de determinação de erros e sua representação. Os modelos de Rauscher [19] e Khoshelham [11] apresentam o erro de maneira semelhante à apresentada em nossa proposta, por isso é útil realizar uma comparação que permita validar os resultados obtidos com nosso método.

O diagrama de dispersão da Figura 5 mostra que os valores de erro do modelo de Rauscher e o nosso variam linearmente, de modo que a quantificação do seu grau de relacionamento pode ser calculada pelo coeficiente de correlação. O coeficiente obtido é 0,99, isto implica que as duas curvas têm uma forte relação positiva. Portanto, a curva obtida com a nossa proposta representa corretamente o erro de profundidade do MS Kinect v1 e o método desta proposta pode ser utilizado satisfatoriamente em outros sensores 3D.

A proposta de Khoshelham considera uma representação do desvio padrão do erro de profundidade como uma função da distância ao padrão, para que uma comparação apropriada com nossos resultados possa ser feita o desvio padrão do erro RMS em cada distância e calculado. O coeficiente de correlação entre as duas curvas de desvio padrão do erro é 0,98.

VI. CONCLUSÃO

Neste trabalho, propomos um método genérico para a estimação do erro RMS nos dados de profundidade fornecidos por qualquer dispositivo capaz de capturar uma imagem RGB e gerar um mapa de profundidade de uma mesma cena, ao mesmo tempo. Uma série de cantos extraídos de um tabuleiro de xadrez são usadas para encontrar uma nuvem de pontos estimada que é comparada com outra nuvem chamada de ideal, representando os valores verdadeiros (*ground truth*). Usando

inicialmente um procedimento de registro 3D, a estimação do erro RMS é realizada depois.

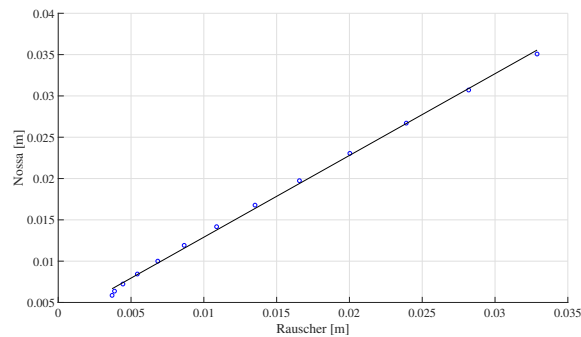
Dos experimentos feitos, conclui-se que a estimação do erro RMS nos dados de profundidade da ZED está diretamente relacionada à resolução da câmera. Quando esta resolução é baixa, a distância de captura de profundidade é pequena e o erro RMS é maior. Observa-se que em todas as resoluções analisadas, a partir de 15 metros, a câmera ZED começa a fornecer alguns dados de profundidade *NaN* para determinados setores do padrão. A falta de dados de profundidade se reflete diretamente do fato de que a grandes distâncias entre o padrão e a câmera, o valor do erro RMS vai rapidamente para o infinito. De fato, em uma breve análise a partir dessas experiências, concluímos que a câmera ZED pode ser usada com boa exatidão em uma distância de até 16 metros. Em distâncias superiores a esta, as curvas mostram erros que podem não ser aceitáveis dependendo da aplicação. Para a robótica terrestre, por exemplo, isso seria considerado uma distância de trabalho muito boa, no entanto, para a robótica aérea, essa distância pode precisar ser aumentada principalmente para robôs aéreos de asa fixa.

Da comparação entre os sensores 3D, pode-se dizer que para aplicações nas quais se pretende utilizar dados de profundidade de até 2 m é preferível usar o MS Kinect v2, pois seu erro RMS é menor do que a ZED e o Kinect v1. Para distâncias superiores a esta pode-se usar a ZED para obter aplicações com o menor erro RMS, porém com alto custo de processamento. O MS Kinect v2 com um mapa de profundidade de 512×424 px tem menor erro RMS do que o Kinect v1, até os 3 m, superior a essa distância, seu erro é maior do que na ZED.

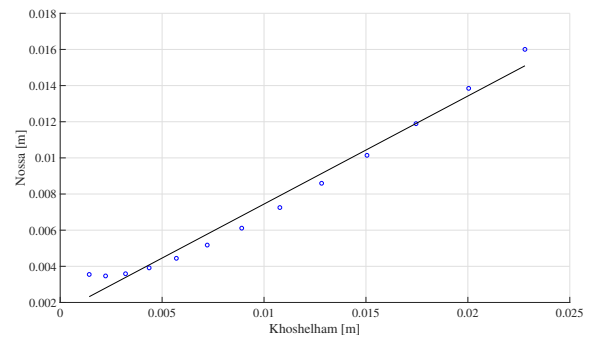
A partir dos experimentos feitos, nota-se que o método proposto é eficaz, propiciando estimar de forma correta o erro RMS nos dados de profundidade. Assim, este trabalho deverá se tornar uma contribuição importante para outros trabalhos futuros que necessitem, como parâmetro útil, da exatidão dos dados de profundidade fornecidos pelos sensores 3D.

AGRADECIMENTOS

Agradecemos ao CNPq e à CAPES, Agências Brasileiras de Financiamento à Pesquisa, pelas bolsas de pesquisa de Luis



(a) Rauscher vs. Nossa proposta (Erro RMS).



(b) Khoshelham vs. Nossa proposta (σ do erro RMS).

Figura 5. Gráficos de dispersão da comparativa dos erros de profundidade do Kinect v1.

E. Ortiz (CAPES), Elizabeth V. Cabrera (CAPES) e Luiz M. G. Gonçalves (CNPq).

PUBLICAÇÕES

Como resultado do trabalho de conclusão do Mestrado de Luis Enrique Ortiz Fernandez, três artigos foram publicados efetivamente (um em anais de eventos e os outros em periódicos): o primeiro no SIBGRAP'16 *Workshop of Works in Progress*, disponível em: <http://urlib.net/8JMKD3MGPAW/3MAK52E>, o segundo no periódico ELCVIA (*Electronic Letters on Computer Vision and Image Analysis*), estrato B2 pelo Qualis CAPES, disponível em: <http://elcvia.cvc.uab.es/article/view/v17-n1-ortiz>, e o terceiro no periódico Sensors (JCR de 2,47, e A1 pelo Qualis CAPES), disponível em: <https://www.mdpi.com/1424-8220/18/9/3122>, com DOI 10.3390/s18093122.

REFERÊNCIAS

- [1] W. Zhao and N. Nandhakumar, "Effects of camera alignment errors on stereoscopic depth estimates," *Pattern Recognition*, vol. 29, no. 12, pp. 2115–2126, 1996.
- [2] S. D. Blostein and T. S. Huang, "Error Analysis in Stereo Determination of 3-D Point Positions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-9, no. 6, pp. 752–765, 1987.
- [3] C. Chang and S. Chatterjee, "Quantization error analysis in stereo vision," in *Conference Record of the Twenty-Sixth Asilomar Conference on Signals, Systems Computers*, 1992, pp. 1037–1041 vol.2.
- [4] Y. Xiong and L. Matthies, "Error analysis of a real-time stereo system," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1997, pp. 1087–1093.
- [5] Z. Yang and Y. Wang, "Error analysis of 3D shape construction from structured lighting," *Pattern Recognition*, vol. 29, no. 2, pp. 189–206, 1996.
- [6] A. Maimone and H. Fuchs, "Reducing interference between multiple structured light depth sensors using motion," in *IEEE Virtual Reality Workshops (VRW)*, 2012, pp. 51–54.
- [7] G. Sansoni, M. Carocci, and R. Rodella, "Calibration and performance evaluation of a 3-D imaging sensor based on the projection of structured light," *IEEE Transactions on Instrumentation and Measurement*, vol. 49, no. 3, pp. 628–636, 2000.
- [8] Y. Kuzuhara, K. Kawamura, R. Yoshitoshi, T. Tamaki, S. Sugai, M. Ikegami, Y. Kurokawa, T. Obitsu, M. Okita, T. Sugino, and T. Yasuda, "A preliminary study for predicting body weight and milk properties in lactating Holstein cows using a three-dimensional camera system," *Computers and Electronics in Agriculture*, vol. 111, pp. 186–193, 2015.
- [9] B. Galna, G. Barry, D. Jackson, D. Mhiripiri, P. Olivier, and L. Rochester, "Accuracy of the Microsoft Kinect sensor for measuring movement in people with Parkinson's disease," *Gait & Posture*, vol. 39, no. 4, pp. 1062–1068, 2014.
- [10] S. Cloix, G. Bologna, V. Weiss, T. Pun, and D. Hasler, "Low-power depth-based descending stair detection for smart assistive devices," *EURASIP Journal on Image and Video Processing*, vol. 2016, no. 1, p. 33, 2016.
- [11] K. Khoshelham, "Accuracy Analysis of Kinect Depth Data," *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 3812, pp. 133–138, 2011.
- [12] H. Haggag, M. Hossny, D. Filippidis, D. Creighton, S. Nahavandi, and V. Puri, "Measuring depth accuracy in RGBD cameras," in *Signal Processing and Communication Systems (ICSPCS), 7th International Conference*, 2013, pp. 1–7.
- [13] L. Yang, L. Zhang, H. Dong, A. Alelaiwi, and A. E. Saddik, "Evaluating and Improving the Depth Accuracy of Kinect for Windows v2," *IEEE Sensors Journal*, vol. 15, no. 8, pp. 4275–4285, 2015.
- [14] J. Aguilar, F. Torres, and M. Lope, "Stereo vision for 3D measurement: accuracy analysis, calibration and industrial applications," *Measurement*, vol. 18, no. 4, pp. 193–200, 1996.
- [15] P. Luo and S. Liou, "Measurement of curved surface by stereo vision and error analysis," *Optics and Lasers in Engineering*, vol. 30, no. 6, pp. 471–486, 1998.
- [16] C. Chang, S. Chatterjee, and P. R. Kube, "A quantization error analysis for convergent stereo," in *Proceedings of 1st International Conference on Image Processing*, vol. 2, 1994, pp. 735–739 vol.2.
- [17] M. Kytö, M. Nuutinen, and P. Oittinen, "Method for measuring stereo camera depth accuracy based on stereoscopic vision," *Proc. SPIE*, vol. 7864, pp. 78 640I–78 640I–9, 2011.
- [18] S. Hussmann, T. Ringbeck, and B. Hagebeuker, "A performance review of 3D TOF vision systems in comparison to stereo vision systems," in *Stereo vision*. InTech, 2008.
- [19] G. Rauscher, D. Dube, and A. Zell, "A Comparison of 3D Sensors for Wheeled Mobile Robots," in *Intelligent Autonomous Systems 13*, E. Menegatti, N. Michael, K. Berns, and H. Yamaguchi, Eds. Cham: Springer International Publishing, 2016, pp. 29–41.
- [20] A. Geiger, F. Moosmann, O. Car, and B. Schuster, "Automatic camera and range sensor calibration using a single shot," in *2012 IEEE International Conference on Robotics and Automation*, May 2012, pp. 3936–3943.
- [21] K. S. Arun, T. S. Huang, and S. D. Blostein, "Least-Squares Fitting of Two 3-D Point Sets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-9, no. 5, pp. 698–700, 1987.
- [22] R. S. Esfandiari, *Numerical Methods for Engineers and Scientists Using MATLAB®*. CRC Press, 2017.
- [23] L. Ortiz, E. Cabrera, and L. Gonçalves, "3DRT – Sistema Embarcado para Reconstrução 3D em Tempo Real," in *Workshop of Works in Progress (WIP) in the 29th Conference on Graphics, Patterns and Images (SIBGRAP'16)*, São José dos Campos, SP, Brazil, 2016.
- [24] L. E. Ortiz, E. V. Cabrera, and L. M. Gonçalves, "Depth Data Error Modeling of the ZED 3D Vision Sensor from Stereolabs," *ELCVIA: Electronic Letters on Computer Vision and Image Analysis*, vol. 17, no. 1, pp. 0001–15, 2018.