**Exposure optimization for warming of shapes in the automotive industry**

Královcová, J.

2010

**Institute of Computer Science**
**Academy of Sciences of the Czech Republic**

# Exposure optimization for warming of shapes in the automotive industry

Jiřina Královcová, Ladislav Lukšan, J Mlýnek

Technical report No. 1080

September 2010

Pod Vodárenskou věží 2, 182 07 Prague 8 phone: +420 2 688 42 44, fax: +420 2 858 57 89,
e-mail:e-mail:ics@cs.cas.cz

Institute of Computer Science
Academy of Sciences of the Czech Republic

# Exposure optimization for warming of shapes in the automotive industry

Jiřina Královcová, Ladislav Lukšan, J Mlýnek [1]

Technical report No. 1080

September 2010

Abstrakt:

This contribution contains a description and comparison of two methods applied to exposure optimization for warming the shapes in the automotive industry.

Keywords:

[1]Tato práce byla vytvořena v rámci centra excelence MŠMT 1M0554 a podpořena výzkumným záměrem AV0Z10300504 AVČR.

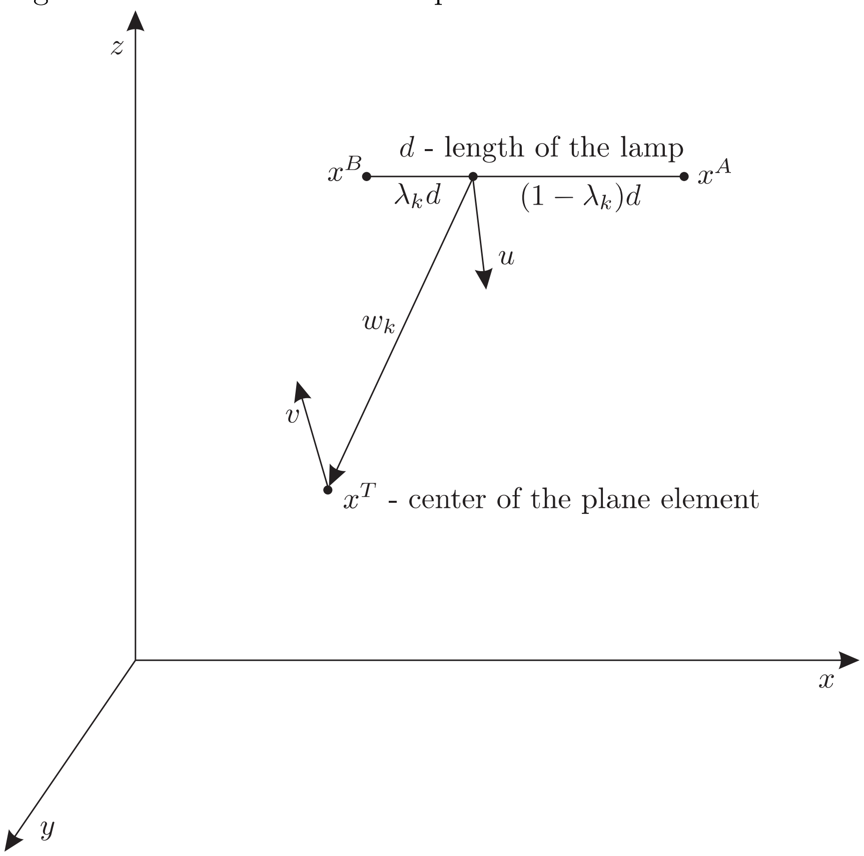
# 1 Introduction

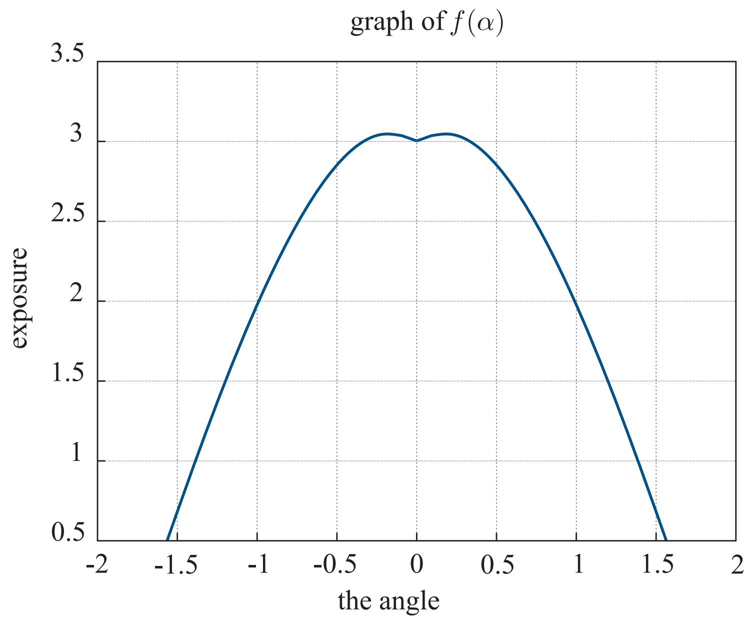
Consider an aluminium shape weighting approximately 300 kg. This shape should be uniformly warming by using approximately 100 lamps of the same performance to reach $270^o C$ temperature.

Every lamp is defined by the coordinates of its side points $A$, $B$ and the lighting direction $u$ (9 parameters). The length $d$ of all lamps is the same.

The shape surface is defined by using approximately 10000 plane elements. Every plane element is represented by the coordinates of its center $T$ and its outer normal $d$ of unit length (6 parameters).

The initial coordinates of all lamps are given. We require a uniform exposure of the shape by seeking a suitable choice of the lamp coordinates.






$$f(\alpha) = 3.0\cos(\alpha) + 0.5\,|\sin(\alpha)|$$

# 2 Formulation of optimization problem with constraints.

## 2.1 Equations for the exposure of a plane element by a lamp.

Let $x^T = (x_1^T, x_2^T, x_3^T)$ be the center of the plane element, $x^N = (x_1^N, x_2^N, x_3^N)$ be the normal of the plane element, $x^A = (x_1^A, x_2^A, x_3^A)$, $x^B = (x_1^B, x_2^B, x_3^B)$ be side points of the lamp and $x^S = (x_1^S, x_2^S, x_3^S)$ be the lighting direction of the lamp. We also denote $v = -x^N$, $u = x^S$ and use the following constraints

$$\sum_{i=1}^{3} (x_i^S)^2 = 1,$$

$$\sum_{i=1}^{3} x_i^S (x_i^B - x_i^A) = 0,$$

$$\sum_{i=1}^{3} (x_i^B - x_i^A)^2 = d^2,$$

where $d$ is the length of the lamp. The first constraint ensures the unit length of vector $x^S$, the second its orthogonality to the axis of the lamp, and the third stabilize the length of the lamp.

The lamp is a linear body of the length $d$, consisting of $p$ lighting elements of lengths $d_k = d/(p-1)$, $1 < k < p$. The distance between the lighting element and the center of the plane element is expressed as

$$w_k = x^T - (1-\lambda_k)x^A - \lambda_k x^B, \quad \lambda_k = \frac{k-1}{p-1},$$

where $1 \le k \le p$. The exposure $I$ of the given plane element by the given lamp is given by the formula

$$I = \sum_{k=1}^{p} I_k, \quad I_k = \left( 3\alpha_k + \frac{1}{2}\sqrt{1-\alpha_k^2} \right) \frac{\beta_k}{\|w_k\|^2} d_k,$$

where

$$\alpha_k = \frac{u^T w_k}{\|u\| \|w_k\|} = \tilde{u}^T \tilde{w}_k, \quad \beta_k = \frac{v^T w_k}{\|v\| \|w_k\|} = \tilde{v}^T \tilde{w}_k,$$

and

$$\tilde{u} = u/\|u\|, \quad \tilde{v} = v/\|v\|, \quad \tilde{w_k} = w_k/\|w_k\|.$$

Analytical expression of derivatives of the exposure $I$ with respect to the elements of vectors $x^A$, $x^B$, $x^S$ (elements of vectors $x^T$, $x^N$ are constants, since the shape surface is invariant) has the form

$$\begin{aligned}
\frac{\partial I}{\partial x_i^A} &= \sum_{k=1}^{p} \frac{\partial I_k}{\partial x_i^A} = -\sum_{k=1}^{p} (1-\lambda_k) \frac{\partial I_k}{\partial w_{ik}}, \\
\frac{\partial I}{\partial x_i^B} &= \sum_{k=1}^{p} \frac{\partial I_k}{\partial x_i^B} = -\sum_{k=1}^{p} \lambda_k \frac{\partial I_k}{\partial w_{ik}} \\
\frac{\partial I}{\partial x_i^S} &= \sum_{k=1}^{p} \frac{\partial I_k}{\partial x_i^S} = \sum_{k=1}^{p} \frac{\partial I_k}{\partial u_i},
\end{aligned}$$

where

$$\begin{aligned}
\frac{\partial I_k}{\partial u_i} &= \left( 3 - \frac{1}{2} \frac{\alpha_k}{\sqrt{1-\alpha_k^2}} \right) \frac{\beta_k}{\|w_k\|^2} \frac{\partial \alpha_k}{\partial u_i} d_k, \\
\frac{\partial I_k}{\partial w_{ik}} &= \left( 3 - \frac{1}{2} \frac{\alpha_k}{\sqrt{1-\alpha_k^2}} \right) \frac{\beta_k}{\|w_k\|^2} \frac{\partial \alpha_k}{\partial w_{ik}} d_k \\
&+ \left( 3\alpha_k + \frac{1}{2}\sqrt{1-\alpha_k^2} \right) \left( \frac{1}{\|w_k\|^2} \frac{\partial \beta_k}{\partial w_{ik}} - 2 \frac{\beta_k}{\|w_k\|^4} w_{ik} \right) d_k.
\end{aligned}$$

Furthermore, one has

$$\begin{aligned}
\frac{\partial \alpha_k}{\partial u_i} &= \frac{w_{ik}}{\|u\| \|w_k\|} - \frac{u^T w_k}{\|u\| \|w_k\|} \frac{u_i}{\|u\|^2} = \frac{1}{\|u\|} (\tilde{w}_{ik} - \alpha_k \tilde{u}_i), \\
\frac{\partial \alpha_k}{\partial w_{ik}} &= \frac{u_i}{\|u\| \|w_k\|} - \frac{u^T w_k}{\|u\| \|w_k\|} \frac{w_{ik}}{\|w_k\|^2} = \frac{1}{\|w_k\|} (\tilde{u}_i - \alpha_k \tilde{w}_{ik}), \\
\frac{\partial \beta_k}{\partial w_{ik}} &= \frac{v_i}{\|v\| \|w_k\|} - \frac{v^T w_k}{\|v\| \|w_k\|} \frac{w_{ik}}{\|w_k\|^2} = \frac{1}{\|w_k\|} (\tilde{v}_i - \beta_k \tilde{w}_{ki}),
\end{aligned}$$

and after substitution we obtain

$$\frac{\partial I_k}{\partial u_i} = \left(3 - \frac{1}{2}\frac{\alpha_k}{\sqrt{1-\alpha_k^2}}\right)\frac{\beta_k d_k}{\|u\|\|w_k\|^2}(\tilde{w}_{ik} - \alpha_k \tilde{u}_i)$$

$$\frac{\partial I_k}{\partial w_{ik}} = \left(3 - \frac{1}{2}\frac{\alpha_k}{\sqrt{1-\alpha_k^2}}\right)\frac{\beta_k d_k}{\|w_k\|^3}(\tilde{u}_i - \alpha_k \tilde{w}_{ik})$$
$$+ \left(3\alpha_k + \frac{1}{2}\sqrt{1-\alpha_k^2}\right)\frac{d_k}{\|w_k\|^3}(\tilde{v}_i - 3\beta_k \tilde{w}_{ik}).$$

It is not necessary to known elements of vectors $u$, $v$ and $w_k$, $1 \le k \le p$. We use only their Euclidean norms and elements of the normalized vectors $\tilde{u}$, $\tilde{v}$ and $\tilde{w}_k$, $1 \le k \le p$. Therefore, it is advantageous to normalize vectors $u$, $v$ and $w_k$, $1 \le k \le p$, beforehand.

## 2.2 Objective function and constraints for the uniform exposure.

We have $n_e$ plane elements and $n_l$ lamps. Every plane element can be exposed by several lamps. Let $L_j$ be a set of indices of lamps that expose $j$-th plane element. Choose $1 \le j \le n_e$ and $l \in L_j$. If we denote $I_{jl}$ the exposure of $j$-th element by $l$-th lamp, (this value correspond to the value $I$ from the previous subsection), then the total exposure $I_j$ of $j$-th element is given by the formula

$$I_j = \sum_{l \in L_j} I_{jl}.$$

The derivatives of $I_j$ are computed by the formulas

$$\frac{\partial I_j}{\partial x_{il}^A} = \frac{\partial I_{jl}}{\partial x_{il}^A}, \quad \frac{\partial I_j}{\partial x_{il}^B} = \frac{\partial I_{jl}}{\partial x_{il}^B}, \quad \frac{\partial I_j}{\partial x_{il}^S} = \frac{\partial I_{jl}}{\partial x_{il}^S}, \quad l \in L_j,$$

$$\frac{\partial I_j}{\partial x_{il}^A} = 0, \qquad \frac{\partial I_j}{\partial x_{il}^A} = 0, \qquad \frac{\partial I_j}{\partial x_{il}^A} = 0, \qquad l \notin L_j,$$

where we substitute the previously defined quantities. Let $\overline{I}$ be the prescribed value of the exposure (the same for all elements of the shape surface). Then

$$F(x) = \frac{1}{2}\sum_{j=1}^{n_e}(I_j - \overline{I})^2,$$

where vector $x$ has elements $x_{1l}^A$, $x_{2l}^A$, $x_{3l}^A$, $x_{1l}^B$, $x_{2l}^B$, $x_{3l}^B$, $x_{1l}^S$, $x_{2l}^S$, $x_{3l}^S$, $1 \le l \le n_l$ (nine for every lamp). One has

$$\begin{aligned}
\frac{\partial F(x)}{\partial x_{il}^A} &= \sum_{j=1}^{n_e} (I_j - \overline{I}) \frac{\partial I_j}{\partial x_{il}^A}, \\
\frac{\partial F(x)}{\partial x_{il}^B} &= \sum_{j=1}^{n_e} (I_j - \overline{I}) \frac{\partial I_j}{\partial x_{il}^B}, \\
\frac{\partial F(x)}{\partial x_{il}^S} &= \sum_{j=1}^{n_e} (I_j - \overline{I}) \frac{\partial I_j}{\partial x_{il}^S},
\end{aligned}$$

where we substitute quantities computed in the previous relations. The prescribed value of the exposure is determined initially using the formula

$$\overline{I} = \frac{1}{n_e} \sum_{j=1}^{n_e} I_j.$$

The objective function $F(x)$ is minimized on the feasible region given by the equality constraints

$$\begin{aligned}
c_{1l}(x) &= \sum_{i=1}^{3} (x_{il}^S)^2 = 1, \\
c_{2l}(x) &= \sum_{i=1}^{3} x_{il}^S (x_{il}^B - x_{il}^A) = 0, \\
c_{3l}(x) &= \sum_{i=1}^{3} (x_{il}^B - x_{il}^A)^2 = d^2,
\end{aligned}$$

where $1 \le l \le n_l$ (three for every lamp). It holds

$$\begin{aligned}
&\frac{\partial c_{1l}(x)}{\partial x_{il}^A} = 0, && \frac{\partial c_{1l}(x)}{\partial x_{il}^B} = 0, && \frac{\partial c_{1l}(x)}{\partial x_{il}^S} = 2x_{il}^S, \\
&\frac{\partial c_{2l}(x)}{\partial x_{il}^A} = -x_{il}^S, && \frac{\partial c_{2l}(x)}{\partial x_{il}^B} = x_{il}^S, && \frac{\partial c_{2l}(x)}{\partial x_{il}^S} = x_{il}^B - x_{il}^A, \\
&\frac{\partial c_{2l}(x)}{\partial x_{il}^A} = -2(x_{il}^B - x_{il}^A), && \frac{\partial c_{2l}(x)}{\partial x_{il}^B} = 2(x_{il}^B - x_{il}^A), && \frac{\partial c_{2l}(x)}{\partial x_{il}^S} = 0
\end{aligned}$$

and remaining derivatives are zeroes. The constraints are sparse, so the memory size and the number of arithmetic operations are not large.

The described problem consists in the minimization of a sum of squares with respect to nonlinear equality constraints. The number of partial functions in the sum of squares is

$n_e \sim 10000$ (the number of the plane elements). The number of variables is $9n_l \sim 900$ (nine for every lamp). The Hessian matrix of the objective function is not sparse. The number of nonlinear equality constraints is $3n_l \sim 300$ (three for every lamp). The Jacobian matrix of nonlinear equality constraints is sparse. These facts have an influence to the choice of the numerical method.

## 2.3 The recursive quadratic programming method.

We want to find a local minimum of the twice continuously differentiable function $F : R^n \to R$, on the feasible set given by the equality constraints

$$c_i(x) = 0, \qquad 1 \le i \le m.$$

Here $x \in R^n$ a $c_i : R^n \to R$, $1 \le i \le m \le n$, are twice continuously differentiable functions. If the LICQ constraint qualification (the linear independence of gradients of the constraint functions) is satisfied, the necessary conditions for the local minimum have the form

$$\begin{aligned} \nabla F(x) + A(x)u &= 0, \\ c(x) &= 0. \end{aligned}$$

This is a set of $n + m$ nonlinear equations for unknown vectors $x \in R^n$ and $u \in R^m$, where $A(x)$ is the Jacobian matrix of the mapping $c(x)$ and $u$ is the vector of Lagrange multipliers.

The principle of the recursive quadratic programming method consists in the application of the Newton method to the system of nonlinear equations specifying the necessary conditions for the local minimum. The iterative step of the Newton method has the form

$$\begin{aligned} x_{k+1} &= x_k + \alpha_k d_k^x, \\ u_{k+1} &= u_k + \alpha_k d_k^u, \end{aligned}$$

where $d_k^x$, $d_k^u$ are direction vectors obtained as a solutions of the system of linear equations

$$\begin{bmatrix} G(x_k, u_k) & A(x_k) \\ A(x_k)^T & 0 \end{bmatrix} \begin{bmatrix} d_k^x \\ d_k^u \end{bmatrix} = - \begin{bmatrix} g(x_k, u_k) \\ c(x_k) \end{bmatrix}$$

(the linear KKT system) and $\alpha_k > 0$ is a selected stepsize. Here

$$g(x, u) = \nabla F(x) + \sum_{i=1}^{m} u_i \nabla c_i(x), \quad G(x, u) = \nabla^2 F(x) + \sum_{i=1}^{m} u_i \nabla^2 c_i(x)$$

are the gradient and the Hessian matrix of the Lagrangian function. The exact Hessian matrix $G(x_k, u_k)$ is replaced by its approximation $B_k$ obtained by the BFGS quasi-Newton method. Then

$$\begin{bmatrix} B_k & A_k \\ A_k^T & 0 \end{bmatrix} \begin{bmatrix} d_k^x \\ d_k^u \end{bmatrix} = - \begin{bmatrix} g(x_k, u_k) \\ c(x_k) \end{bmatrix},$$

where $B_1 = I$ ($I$ is the unit matrix of order $n$) and

$$B_{k+1} = B_k + \frac{1}{s_k^T y_k} y_k y_k^T - \frac{1}{s_k^T B_k s_k}(B_k s_k y_k^T + y_k s_k^T B_k),$$

where $s_k = x_{k+1} - x_k = \alpha_k d_k$ and $y_k = g(x_{k+1}, u_{k+1}) - g(x_k, u_{k+1})$.

As a merit function for stepsize selection, the augmented Lagrangian function

$$P_k(\alpha) = F(x_k + \alpha d_k^x) + (u_k + d_k^u)^T c(x_k + \alpha d_k^x) + \frac{\sigma}{2}\|c(x_k + \alpha d_k^x)\|^2$$

is used where $\sigma \geq 0$ is a penalty parameter. If the linear KKT system is solved in such a way that

$$\|G_k d_k^x + A_k d_k^u + g_k\| \leq \overline{\omega}_k \|g_k\| \quad \|A_k^T d_k^x + c_k\| \leq \overline{\omega}_k \|c_k\|,$$

where $0 < \overline{\omega}_k < 1$, then $P_k'(0) < 0$ holds and function $P_k(\alpha)$ is decreasing in the direction $d_k^x$. In this case, we can chose the sepsize in such a way that $\alpha_k = \beta^{j-1}\max(1, \overline{\Delta}/\|d_k^x\|)$, where $\overline{\Delta}$ is the maximum stepsize, $0 < \beta < 1$ is the reduction coefficient and $j \in N$ is the minimum integer such that

$$P_k(\alpha_k) - P_k(0) \leq \varepsilon_1 \alpha_k P_k'(0),$$

where $0 < \varepsilon_1 < 1/2$ is the Armijo parameter. The values $\overline{\Delta} = 1000$, $\beta = 0.5$ and $\varepsilon_1 = 0.0001$ are usually used.

## 2.4 Solving the linear KKT system.

The linear KKT system can be written in the form

$$Kd = \begin{bmatrix} B & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} d^x \\ d^u \end{bmatrix} = \begin{bmatrix} b^x \\ b^u \end{bmatrix} = b.$$

This symmetric system of linear equations, whose matrix is indefinite, is solved by the preconditioned conjugate gradient method with preconditioner

$$C = \begin{bmatrix} D & A \\ A^T & 0 \end{bmatrix},$$

where $D$ is a positive definite diagonal matrix approximating in some sense the main diagonal of $B$. The multiplication of vector $r$ by the matrix $C^{-1}$ can be expressed in the form

$$C^{-1}r = \begin{bmatrix} D^{-1}(r^x - At^u) \\ t^u \end{bmatrix}, \quad t^u = (A^T D^{-1} A)^{-1}(A^T D^{-1} r^x - r^u).$$

**Algorithm:** Set $d_1 = 0$, $r_1 = b$,

$$t_1^u = (A^T D^{-1} A)^{-1}(A^T D^{-1} r_1^x - r_1^u), \quad t_1^x = D^{-1}(r_1^x - At_1^u)$$

and $p_1 = t_1$. For $i \geq 1$ the following steps are performed. If $\|r_i^x\| \leq \overline{\omega}\|b^x\|$ and $\|r_i^u\| \leq \overline{\omega}\|b^u\|$, where $\overline{\omega}$ is a given precision, then set $d = d_i$ and terminate the computation. In the opposite case compute

$$\begin{aligned} q_i &= Kp_i, & \alpha_i &= r_i^T t_i / p_i^T q_i, \\ d_{i+1} &= d_i + \alpha_i p_i, & r_{i+1} &= r_i - \alpha_i q_i, \\ t_{i+1}^u &= (A^T D^{-1} A)^{-1} (A^T D^{-1} r_{i+1}^x - r_{i+1}^u), & & \\ t_{i+1}^x &= D^{-1}(r_{i+1}^x - A t_{i+1}^u), & & \\ \beta_i &= r_{i+1}^T t_{i+1} / r_i^T t_i, & p_{i+1} &= t_{i+1} + \beta_i p_i \end{aligned}$$

and increase $i$ by 1.
Matrix $(A^T D^{-1} A)^{-1}$ need not be computed explicitly, we use its Choleski decomposition $LL^T = A^T D^{-1} A$, where $L$ is a lower triangular matrix.

# 3 Formulation of optimization problem with constraints.

## 3.1 Equations for the exposure of a plane element by a lamp.

Let $x^T = (x_1^T, x_2^T, x_3^T)$ be the center of the plane element, $x^N = (x_1^N, x_2^N, x_3^N)$ be the normal of the plane element, $x^A = (x_1^A, x_2^A, x_3^A)$, $x^B = (x_1^B, x_2^B, x_3^B)$ be side points of the lamp and $x^S = (x_1^S, x_2^S, x_3^S)$ be the lighting direction of the lamp. We also denote $v = -x^N$, $x = x^A$, $u = x^S$. Let $y$ be a vector parallel to the vector $x^B - x^A$, so $x^B - x^A = (y/\|y\|)d$, where $d = \|x^B - x^A\|$.

We assume that the lighting direction of the lamp is mostly perpendicular to the plane of the shape, so the angle between vector $x^S$, which is perpendicular to the vector $y$, and the normal $e = (0, 0, -1)$ of the plane (assumed to be horizontal) is minimal. If the norm of vector $u$ is unit, it can be uniquely determined from vectors $y$ and $e$.

**Věta 1** *Vector*

$$u = \frac{e + \lambda y}{\sqrt{e^T(e + \lambda y)}}, \qquad \lambda = -\frac{e^T y}{y^T y}.$$

*is the solution of the optimization problem*

$$e^T u \ \rightarrow \ \max,$$

$$y^T u = 0,$$
$$u^T u = 1,$$

Since the length of vector $u$ can be arbitrary, we put

$$u = e - \frac{e^T y}{y^T y} y = \tilde{e} - \tilde{e}^T \tilde{y} \tilde{y},$$

where $\tilde{e} = e/\|e\|$ and $\tilde{y} = y/\|y\|$ (vector $e = (0, 0, -1)$ has the unit norm). To compute the gradient of the objective function, we will use the Jacobian matrix $\nabla_y u$ of vector $u$ depending on elements of vector $y$).

**Věta 2** *One has*

$$\nabla_y u = \left(2\,\frac{y\,y^T}{y^T y} - I\right)\frac{e^T y}{y^T y} - \frac{e\,y^T}{y^T y} = \frac{1}{\|y\|}\left((2\,\tilde{y}\tilde{y}^T - I)\,\tilde{e}^T\tilde{y} - \tilde{e}\,\tilde{y}^T\right)$$

The lamp is a linear body of the length $d$, consisting from $p$ lighting elements of lengths $d_k = d/(p-1)$, $1 < k < p$. The distances between the lighting element and the center of the plane element is expressed as

$$w_k = x^T - (1-\lambda_k)x^A - \lambda_k x^B = x^T - x^A - \lambda_k d\frac{y}{\|y\|}, \quad \lambda_k = \frac{k-1}{p-1},$$

where $1 \le k \le p$. It holds

$$\nabla_y w_k = -\frac{\lambda_k d}{\|y\|}\left(I - \frac{y\,y^T}{y^T y}\right) = -\frac{\lambda_k d}{\|y\|}(I - \tilde{y}\tilde{y}^T),$$

so

$$I = \sum_{k=1}^{p} I_k, \quad I_k = \left(3\alpha_k + \frac{1}{2}\sqrt{1-\alpha_k^2}\right)\frac{\beta_k}{\|w_k\|^2}d_k,$$

where

$$\alpha_k = \frac{u^T w_k}{\|u\|\|w_k\|} = \tilde{u}^T\tilde{w}_k, \quad \beta_k = \frac{v^T w_k}{\|v\|\|w_k\|} = \tilde{v}^T\tilde{w}_k.$$

Analytical expression of derivatives of the exposure $I$ with respect to the elements of vectors $x = x^A$, and $y = x^B - x^A$ (elements of vectors $x^T$, $x^N$ are constants, since the shape surface is invariant) has the form

$$\begin{aligned}
\nabla_x I &= \sum_{k=1}^{p}\nabla_x I_k = -\sum_{k=1}^{p}\nabla_{w_k} I_k,\\
\nabla_y I &= \sum_{k=1}^{p}\nabla_y I_k = \sum_{k=1}^{p}(\nabla_y u\nabla_u I_k + \nabla_y w_k \nabla_{w_k} I_k)
\end{aligned}$$

where

$$\begin{aligned}
\nabla_u I_k &= \left(3 - \frac{1}{2}\frac{\alpha_k}{\sqrt{1-\alpha_k^2}}\right)\frac{\beta_k d_k}{\|w_k\|^2}\nabla_u \alpha_k,\\
\nabla_{w_k} I_k &= \left(3 - \frac{1}{2}\frac{\alpha_k}{\sqrt{1-\alpha_k^2}}\right)\frac{\beta_k d_k}{\|w_k\|^2}\nabla_{w_k}\alpha_k\\
&+ \left(3\alpha_k + \frac{1}{2}\sqrt{1-\alpha_k^2}\right)\left(\frac{d_k}{\|w_k\|^2}\nabla_{w_k}\beta_k - 2\frac{\beta_k d_k}{\|w_k\|^4}w_{ik}\right).
\end{aligned}$$

Furthermore, one has

$$\begin{aligned}
\nabla_u \alpha_k &= \frac{w_k}{\|u\|\|w_k\|} - \frac{u^T w_k}{\|u\|\|w_k\|}\frac{u}{\|u\|^2} = \frac{1}{\|u\|}(\tilde{w}_k - \alpha_k \tilde{u}), \\
\nabla_{w_k} \alpha_k &= \frac{u}{\|u\|\|w_k\|} - \frac{u^T w_k}{\|u\|\|w_k\|}\frac{w_k}{\|w_k\|^2} = \frac{1}{\|w_k\|}(\tilde{u} - \alpha_k \tilde{w}_k), \\
\nabla_{w_k} \beta_k &= \frac{v}{\|v\|\|w_k\|} - \frac{v^T w_k}{\|v\|\|w_k\|}\frac{w_k}{\|w_k\|^2} = \frac{1}{\|w_k\|}(\tilde{v} - \beta_k \tilde{w}_k),
\end{aligned}$$

and after substitution we obtain

$$\begin{aligned}
\nabla_u I_k &= \left(3 - \frac{1}{2}\frac{\alpha_k}{\sqrt{1-\alpha_k^2}}\right)\frac{\beta_k d_k}{\|u\|\|w_k\|^2}(\tilde{w}_k - \alpha_k \tilde{u}) \\
\nabla_{w_k} I_k &= \left(3 - \frac{1}{2}\frac{\alpha_k}{\sqrt{1-\alpha_k^2}}\right)\frac{\beta_k d_k}{\|w_k\|^3}(\tilde{u} - \alpha_k \tilde{w}_k) \\
&+ \left(3\alpha_k + \frac{1}{2}\sqrt{1-\alpha_k^2}\right)\frac{d_k}{\|w_k\|^3}(\tilde{v} - 3\beta_k \tilde{w}_k).
\end{aligned}$$

Note that theorem 2 implies

$$\nabla_y u \nabla_u I_k + \nabla_y w_k \nabla_{w_k} I_k = -\frac{1}{\|y\|}\left(\gamma_e(\nabla_u I_k - 2\gamma_u \tilde{y}) + \gamma_u \tilde{e} + \lambda_k d(\nabla_{w_k} I_k - \gamma_{w_k}\tilde{y})\right),$$

where $\gamma_e = \tilde{y}^T \tilde{e}$, $\gamma_u = \tilde{y}^T \nabla_u I_k$ and $\gamma_{w_k} = \tilde{y}^T \nabla_{w_k} I_k$.

## 3.2 Objective function for the uniform exposure.

We have $n_e$ plane elements and $n_l$ lamps. Every plane element can be exposed by several lamps. Let $L_j$ be a set of indices of lamps that expose $j$-th plane element. Choose $1 \le j \le n_e$ and $l \in L_j$. If we denote $I_{jl}$ the exposure of $j$-th element by $l$-th lamp, (this value correspond to the value $I$ from the previous subsection), then the total exposure $I_j$ of $j$-th element $I_j$ is given by the formula

$$I_j = \sum_{l \in L_j} I_{jl}.$$

The derivatives of $I_j$ are computed by the formulas

$$\begin{aligned}
\nabla_{x_l} I_j &= \nabla_{x_l} I_{jl}, \quad \nabla_{y_l} = \nabla_{y_l} I_{jl}, \quad l \in L_j, \\
\nabla_{x_l} I_j &= 0, \qquad\quad \nabla_{y_l} = 0, \qquad\quad l \notin L_j,
\end{aligned}$$

where we substitute the previously defined quantities. Let $\overline{I}$ be the prescribed value of the exposure (the same for all elements of the shape surface). Then

$$F(x) = \frac{1}{2}\sum_{j=1}^{n_e}(I_j - \overline{I})^2,$$

where vector $x$ has elements $x_{1l}$, $x_{2l}$, $x_{3l}$, $y_{1l}$, $y_{2l}$, $y_{3l}$, $1 \le l \le n_l$ (six for every lamp). One has

$$\nabla_{x_l} F(x) = \sum_{j=1}^{n_e} (I_j(x) - \overline{I}) \nabla_{x_l} I_j(x), \quad \nabla_{y_l} F(x) = \sum_{j=1}^{n_e} (I_j(x) - \overline{I}) \nabla_{y_l} I_j(x),$$

where we substitute derivatives computed in the previous relations. The prescribed value of the exposure is determined initially using the formula

$$\overline{I} = \frac{1}{n_e} \sum_{j=1}^{n_e} I_j.$$

The described problem consists in the minimization of a sum of squares without constraints. The number of partial functions in the sum of squares is $n_e \sim 10000$ (the number of the plane elements). The number of variables is $6n_l \sim 900$ (six for every lamp). The Hessian matrix of the objective function is not sparse. These facts have an influence to the choice of the numerical method.

## 3.3 The combined method for minimizing the sum of squares.

Let

$$F(x) = f^T(x) f(x) = \sum_{k=1}^{m} f_k(x),$$

where $f_k(x)$, $1 \le k \le m$, be a twice continuously differentiable functions. Then the gradient $g(x)$ and the Hessian matrix $G(x)$ of the objective function $F(x)$ can be expressed in the form

$$\begin{aligned} g(x) &= J^T(x) f(x) = \sum_{k=1}^{m} f_k(x) g_k(x) \\ G(x) &= J^T(x) J(x) + C(x) = \sum_{k=1}^{m} g_k(x) g_k^T(x) + \sum_{k=1}^{m} f_k(x) G_k(x). \end{aligned}$$

The direction vector is determined by the trust region method in such a way that

$$s_i = \arg \min_{\|s\| \le \Delta_i} Q_i(s),$$

$$\begin{aligned} x_{i+1} &= x_i, & \rho_i(s_i) &\le 0, \\ x_{i+1} &= x_i + s_i, & \rho_i(s_i) &> 0 \end{aligned}$$

$$\begin{aligned} \underline{\beta} \|s_i\| \le \Delta_{i+1} \le \overline{\beta} \|s_i\|, &\quad \rho_i(s_i) < \underline{\rho}, \\ \Delta_i \le \Delta_{i+1} \le \overline{\Delta}), &\quad \rho_i(s_i) \ge \underline{\rho}, \end{aligned}$$

where

$$Q_i(s) = g_i^T s + \frac{1}{2} s^T B_i s \quad \rho_i(s) = \frac{F(x_i + s) - F(x_i)}{Q_i(s)}$$

and $B_i$ is an approximation of $G(x_i)$. The Gauss–Newton method uses the matrix

$$B_i = J_i^T J_i = \sum_{k=1}^{m} g_k(x_i) g_k^T(x_i).$$

We combine the Gauss–Newton method with the BFGS quasi-Newton method. In this case

$$\begin{aligned} B_{i+1} &= J_{i+1}^T J_{i+1}, & (F_i - F_{i+1})/F_i > \underline{\vartheta}, \\ B_{i+1} &= B_i + \frac{y_i y_i^T}{y_i^T d_i} - \frac{B_i d_i (B_i d_i)^T}{d_i^T B_i d_i}, & (F_i - F_{i+1})/F_i \le \underline{\vartheta}, \end{aligned}$$

where $d_i = x_{i+1} - x_i$, $y_i = g_{i+1} - g_i$ and usually $\underline{\vartheta} = 10^{-4}$. This combined method is superlinearly convergent if it is applied to problems with large residuals.

## 4 Numerical comparison.

The objective function defined in Section 2 was minimized, subject to nonlinear equality constraints, by the recursive quadratic programming method described in [3]. More details can be found in [1]. The objective function defined in Section 3 was minimized by the hybrid method described in [2]. Both these methods are implemented in the universal functional optimization system UFO [4].

The following table contains the results obtained by two mentioned methods applied to the four sample problems.

| | Method with constraints | | | | Method without constraints | | | |
|---|---|---|---|---|---|---|---|---|
| Problem | NIT | NFV | Time | F | NIT | NFV | Time | F |
| L1 | 1111 | 4272 | 12.43 | 26.02 | 46 | 105 | 0.33 | 27.11 |
| L2 | 939 | 3551 | 11.07 | 30.41 | 55 | 123 | 0.39 | 30.02 |
| L3 | 312 | 630 | 3.18 | 12.68 | 99 | 226 | 1.40 | 10.60 |
| L4 | 4282 | 50003 | 141.36 | 1.78* | 64 | 142 | 0.39 | 1.20 |

These results demonstrate that the analytical elimination of constraints considerable increases the efficiency of numerical optimization.

## Reference


[1] Královcová, J., Lukšan, L., Mlýnek, J: Optimalizace osvitu pro tepelný ohřev forem v automobilovém průmyslu. Technical report V-1050, ÚIVT AVČR, Praha 2009.

[2] Lukšan L.: Hybrid methods for large sparse nonlinear least squares. Journal of Optimization Theory and Applications, Vol. 89, 1996, pp.575-595.

[3] Lukšan L., Vlček J.: Indefinitely preconditioned inexact Newton method for large sparse equality constrained nonlinear programming problems. Numerical Linear Algebra with Applications, Vol. 5, 1998, pp.219-247.

[4] Lukšan, L., Tůma, M., Vlcek, J., Ramešová, N., Šiška, M., Hartman, J., Matonoha, C.: UFO 2008 - Interactive system for universal functional optimization. Technical Report V-1151, ÚIVT AVČR, Praha 2011.